ALL DATA ARE HUMAN: THE HUMAN INFRASTRUCTURE OF CIVIC DATA

A Dissertation
Presented to
The Academic Faculty

by

Firaz Ahmed Peer

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in
Digital Media

School of Literature, Media and Communication
Ivan Allen College of Liberal Arts
Georgia Institute of Technology
August 2020

ALL DATA ARE HUMAN: THE HUMAN INFRASTRUCTURE OF CIVIC DATA

Approved by:

Dr. Carl DiSalvo, Advisor
School of Interactive Computing
Georgia Institute of Technology

Dr. Paul Mihailidis
Emerson College

Dr. Chris Le Dantec
School of Literature, Media and Communication
Georgia Institute of Technology

Dr. Yanni Loukissas
School of Literature, Media and Communication
Georgia Institute of Technology

Dr. Nassim Parvin
School of Literature, Media and Communication
Georgia Institute of Technology

Date Approved: July 24, 2020

For Ahlaam

ACKNOWLEDGEMENTS

I recognize the immense amount of luck and privilege it takes to complete a dissertation, especially during a global pandemic and nationwide unrest. This dissertation would not have been possible without the commitment and support of the multiple communities I was a part of, both personal and professional. I would like to thank my advisor, Carl DiSalvo and all my committee members for believing in me and standing by me over the years that it took me to develop as a scholar. Exchanging ideas with them and observing their thought process always energized me to push myself further. My early interactions with Ali Mazalek, Florian 'Floyd' Mueller, Michael Nitsche and Mark Guzdial convinced me that academic research was my true calling.

I also want to thank my friends, colleagues and acquaintances, of whom there are too many to mention, for the collaborations and conversations that sensitized me to the nuances of community based participatory design and helped me articulate my research. The different research labs I participated in afforded me with the time and space to explore multiple research trajectories and hone my skills. The support staff at the Ivan Alan College was always there to ensure that my technology needs were met.

I'm indebted to the Westside neighborhood residents and community leaders who participated and supported me in my research efforts. They were all extremely generous with their time and critical with their feedback, which only helped strengthen my dissertation. I'd also like to thank Jacqueline Royster and her team at the Westside Communities Alliance, Sherry, Mackenzie and Katie for trusting me with their project and supporting me throughout my dissertation research.

I'd like to thank my parents, who always emphasized the importance of education and laid the foundation upon which I could build my intellectual ambitions. I will be forever grateful to my brother Shahab and the rest of my family who supported me unconditionally throughout my PhD, despite not knowing exactly what I was doing or what would come out of it. Friends like Ramik, Vasu, Karthik, Sushma, Scooby, Santhoo, Seema, Divya, Ragu, Prabhu, Aman and Ibrahim became family and made enduring the PhD journey a lot more enjoyable. I'd like to thank my in-laws and my sisters Ala and Hanan, who actively participated and supported my dream of becoming a scholar. My wife Naila endured countless weekdays and weekends of me being immersed in my work. It is thanks to her patience, love and care that I was able to focus on my research, while she took care of everything else. I hope to make it up to her in the coming years. And finally, I want to thank my daughter Ahlaam, whose warmth, smile and energy made every moment of this journey worthwhile, and to whom I dedicate this entire dissertation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

ICT	Information and Communication Technologies
API	Application Programming Interface
CSCW	Computer Supported Collaborative Work and Social Media
CWK	Communities Who Know
FOIA	Freedom of Information Act
GT	Georgia Tech
HCI	Human Computer Interaction
ICTD	Information Communication Technologies and Development
II	Information Infrastructures
NPU	Neighborhood Planning Unit
PD	Participatory Design
PSM	Public Safety Module
STS	Science & Technology Studies
WCA	Westside Communities Alliance

SUMMARY

This dissertation is grounded in issues related to the publicizing of data, which include issues of equitable access, interpretation and use. By engaging with scholarship from Human Computer Interaction and Science and Technology Studies, I contribute to a situated understanding of the local values and infrastructural arrangements that are required to build, use and maintain equitable data infrastructures that would enable marginalized communities to benefit from the publicizing of data through dashboards. I do this by taking a participatory design based anthropological approach in which I collaborate with local community leaders in order to foreground their needs and values when reimagining their civic data infrastructure.

Doing so led me to identify the key elements of the human infrastructure that need to be considered when designing civic data infrastructures with resource-constrained communities. Bringing these elements of the human infrastructure together and reflecting on how my role as a design researcher changed during the scope of this project, I argue that *all data are human, and the way we do justice to them is by identifying and building relationships between the human elements of the civic data infrastructures that we are trying to build.* This implies that we focus on identifying the human actors that are crucial to these civic data infrastructures, strengthen their working relationships and prioritize their values and needs by including them in our infrastructuring efforts. Ultimately, I hope this dissertation helps researchers and practitioners move beyond the mere publicizing of data as a strategy for data equity, but instead think about realigning the human elements of the underlying data infrastructure in order to empower communities.

CHAPTER 1. INTRODUCTION

The power and prevalence of data in our lives has brought attention to the issue of the data divide, where those with time, knowledge, skills and resources to analyze data benefit disproportionately when compared to those from socioeconomically marginalized communities who lack these privileges. Across the US, these marginalized communities are being left behind in the ongoing data revolution, as individuals and institutions with more data, skills, knowledge, and power get to decide what is best for these communities (Jacobson, 2017). The United Nations has also recognized the data divide as a key barrier that needs to be overcome in order to help developing countries achieve their sustainable development goals (Kirkpatrick & Vacarelu, 2018).

The call to bridge this data divide has come from scholars of communication, critical data studies as well as community informatics (Andrejevic, 2014; Boyd & Crawford, 2012; Gurstein, 2003). Some have debated the very idea of open data as a resource (Gurstein, 2011), while others have proposed strategies that might help such marginalized communities use data to support their advocacy needs (Meng & DiSalvo, 2018). Data have also been advocated for as a human right (Latonero & Gold, 2016). One of the strategies that individuals and organizations are resorting to is to make data publicly available so they are easier to access. Such publicly accessible data are being made available on the Internet in a variety of formats and interfaces like API's (Application Programming Interfaces), FOIA (Freedom of Information Act) requests and data dashboards among others.

Data dashboards, which have been important decision-making tools within the public and private sector are now being made for public consumption of data as well. Stephen Few defines dashboards as "a visual display of the most important information needed to achieve one or more objectives consolidated on a single screen, so it can be monitored and understood at a glance" (Few, 2006b). Such data dashboards promise transparency, efficiency and accountability. They hope to offer citizens with the data and insights about how their cities, which can be used to keep city officials accountable for what they have promised. But again, such public data tend to unequally benefit socioeconomically privileged communities who have the skills and resources to access, interpret and make use of this data over communities that are underserved (Erete & Burrell, 2017). The motivation behind my dissertation is to understand what it takes to design, use and maintain data dashboards that would allow marginalized communities to equitably share in the benefits afforded by the publicizing of data through dashboards. In addition to answering the call to bridge the data divide, such research will also offer practical strategies that will allow us to create public data dashboards that serve the data needs of all communities, by catering to those of the marginalized.

Scholarship that addresses the sociotechnical needs of minoritized communities exists in abundance (Dillahunt & Veinot, 2018; Erete et al., 2018; C. A. Le Dantec & Edwards, 2008; Wheeler & Dillahunt, 2018) and so does scholarship on the use of data dashboards within business and government settings (Bartlett & Tkacz., 2017; Biehl et al., 2007; D. Edwards & Thomas, 2005; Few, 2006a; Shen-Hsieh & Schindl, 2002). Few scholars have brought these domains together to examine the potential for data dashboards as an

equitable resource of data for marginalized communities. My dissertation examines the potential of public data dashboards as a resource for equitable data access and use.

The term equity deserves further clarification. Equity implies that the outcomes one receives are directly proportional to the inputs they put in (Wagstaff, 1994). Therefore, in order to determine whether a data resource is equitable or not, one must know what the communities who use the data bring to the table. This knowledge, when coupled with what these communities actually do with the data will help determine if a given data resource is equitable or not, and how it can serve as a resource for justice.

But equity is just one of the many principles that scholars consider when conceptualizing a just distribution of resources (Wagstaff, 1994). Scholars have also considered equality and need as alternate allocation mechanism for justice (Cohen & Greenberg, 1982). Equality holds that all humans are entitled to equal outcomes (Sampson, 1975). Such a stance is problematic because it claims that those who are privileged and those who are not receive the same benefits, irrespective of their contributions, merits, needs or other inherent systemic biases (Kahn et al., 1982). Alternatively, considering only needs as the determining factor for the just distribution of resources is also imprecise, as it doesn't specify what needs are valid or justified. Given that there are a number of different conceptualizations of justice within community settings, one of the motivations in this dissertation is to determine how community leaders conceptualize the values of equity when they use data dashboards. I choose equity over other values of data justice because it was one of the goals with which the dashboard I am studying was setup in the first place.

My scholarship is situated in the City of Atlanta, within a group of communities that are commonly referred to as the Historic Westside. The Historic Westside refers to a collection of neighborhoods in Atlanta, which have a rich socio-cultural history dating back to the American Civil Rights movement and the work of Martin Luther King Jr, for whom these neighborhoods were home. Once a prosperous community with a thriving middle-class population, many of these inner-city neighborhoods were left socially and economically isolated by the white flight that followed the desegregation of schools. The economic downturn of the late nineties led to further urban blight, abandoned homes and foreclosures and the predictable illicit elements that accompany hard times – crime and drugs.

It is within this context that the Dean of the Ivan Alan College of Liberal Arts at Georgia Institute of Technology, a top ranked research university neighboring the Westside, partnered with community leaders in these neighborhoods to create the Westside Communities Alliance Data Dashboard that would offer residents, researchers and organizations the data they needed to advocate for sociopolitical change within their neighborhoods. The dashboard would include a mix of quantitative and qualitative data that would highlight the neighborhoods assets and their historic significance, while also providing data about community issues that were of primary concern, like education, transportation, public safety, water, and environment. Such a dashboard is different from the common perception of dashboards that we see in our cars and web applications. I explore the question of what exactly a dashboard is in the next section and follow with a more detailed framing of the research question I hope to answer through this dissertation.

1.1 What are Data Dashboards

The question of what exactly qualifies as a dashboard has a variety of responses. In their survey of 83 public data dashboards, (Sarikaya et al., 2019) found a broad interpretation of data dashboards which vary from single-screen static pages to multi-screen interactive websites that are being used for decision making, learning, communication and motivation. The kinds of features to incorporate in the dashboard and challenges one faces in doing so are largely driven by the organizational needs. The authors identified these dashboards based on their visual and functional genre and offered 15 distinguishing factors that they categorize between purpose, audience, visual & interactive features, and data semantics. While many of these dashboard examples come from the business domain, they are increasingly being adopted by city governments and non-profit organizations as a way to better engage with their communities.

Indexes and indicators have become the default manner in which cities and nations measure, predict and track performance. In an effort to become "smart" and efficient many city and national governments have built customized numeric data dashboards for its citizens and administrators alike (Goldsmith & Crawford, 2014; Kitchin, 2014). These cover a broad range from transactional and performance dashboards (Perez & Rushing, 2007), to visualizations and community indicators (Baskett et al., 2008). Kitchin, Lauriault, et al. (2015) broadly classify urban dashboards based on their functionality and features under (i) indicators, (ii) benchmarking and (iii) real time dashboards. Indicator dashboards can include single (employment rate, high school graduation) or composite indicators (health index, livability index) that report on the current state of affairs. Georgia's Online Analytical Statistical Information System[1] (Figure 1) supports web-

[1] https://oasis.state.ga.us/

based tools that offer data and insights into specific phenomena like mortality/morbidity, maternal child health, motor vehicle crashes etc., which can be used to diagnose, measure performance towards a target, and even for predictive purposes. An example of a composite indicator dashboard would be the Health, Environment, and Liveability Platform[2] (Figure 2), which uses multiple datasets, indices, maps and graphs to offer information on the relationship between health, demographics, and the built environment for communities in Fulton County.

Figure 1: Georgia's Online Analytical Statistical Information System.

[2] http://geospatial.gatech.edu/PICH/

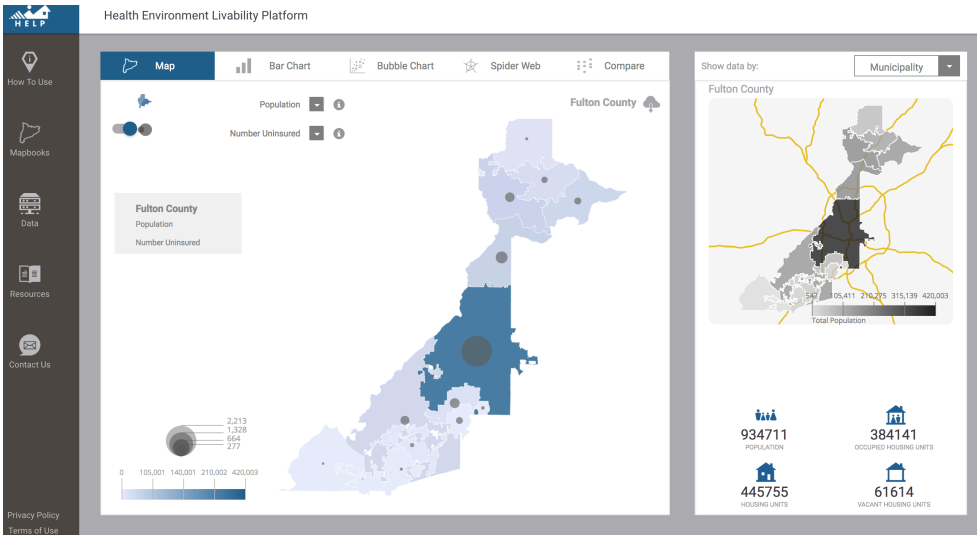


Figure 2: Health, Environment, and Liveability Platform (HELP) developed by the Centre for Geographic Information Systems at Georgia Tech.

Benchmarking dashboards allow users to compare indicators, processes or even policies and the impacts they are having within and across cities. The Fulton County Strategy and Performance Management Office[3] (Figure 3) tracks performance on six priority areas and makes their reports available through their performance dashboard. The Auditor's Office in the City of Atlanta offers a Recommendations Dashboard, which visualizes the number of recommendations the office has received. The Westside Future Fund Data Dashboard[4] (Figure 4) was designed keep investors and residents of the Westside neighborhoods informed of the various developments taking place in their communities, and the impact they were having. The dashboard uses data from the Census, American Community Survey (ACS), Atlanta Police Department, Georgia Department of Education to offer historic trends and track progress across four impact areas, mixed income communities, safety and security, cradle to career education, and

[3] https://performance.fultoncountyga.gov/

[4] http://westsideprogress.org/

community health and wellness. These public facing data dashboards promise access, simplicity and transparency so anyone can benefit from the insights they have to offer.

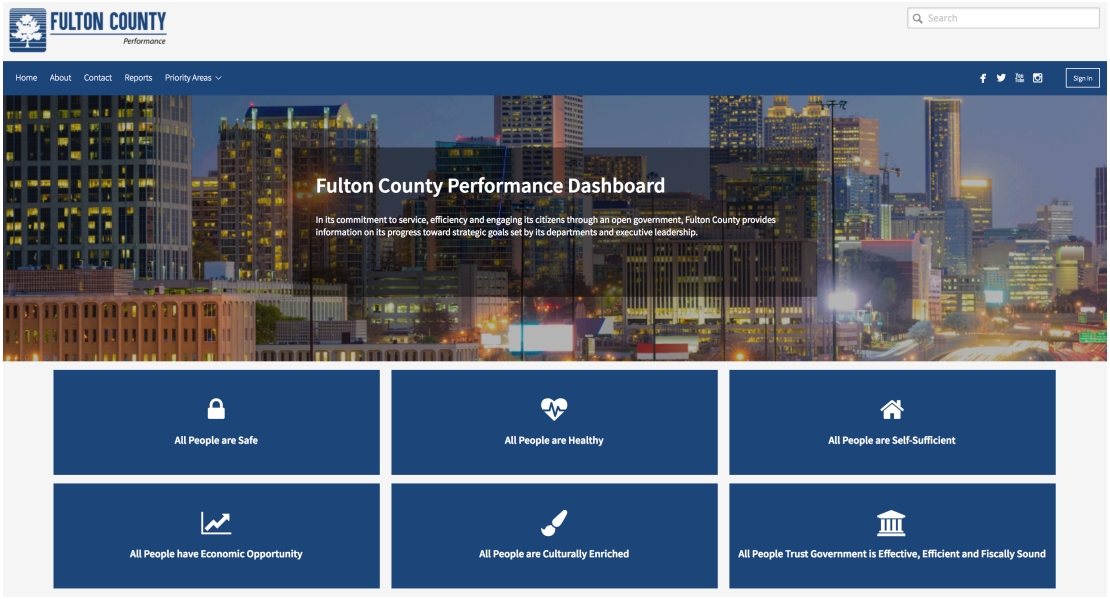


Figure 3: Different dashboards offered by the Fulton County Strategy and Performance Management Office.

Real time dashboards promise up to date and live information about the city. The words "real time" are also interpreted differently based on application and use. Visualizations of traffic and weather tend to be the most common examples of real time data dashboards. The U.S Geological Survey maintains a web interface which allows users to look up real time data about streamflow, water-quality, groundwater levels, precipitation across the United States[5]. Another example comes from Atlanta's Police Department[6] (Figure 5), who update their city-wide crime data on a daily basis. They make this data along with historic data about crime available on their open data website. Outside of Atlanta, the Centro De Operacoes Prefeitura Do Rio[7] (COR) translated as the

[5] https://waterdata.usgs.gov/nwis

[6] http://opendata.atlantapd.org/Default.aspx

[7] http://cor.rio/

Operations Center for the Prefecture of Rio, in Rio de Janeiro, Brazil, the Dublin Dashboard[8] and several others that have been developed under the "Smart Cities" agenda, visualize real time traffic and environmental data along with datasets from around the city. In the lexicon of smart cities, such numeric data dashboards are being touted as a panacea that will make the city administration more transparent and efficient, while simultaneously making the citizens more empowered and civically engaged.

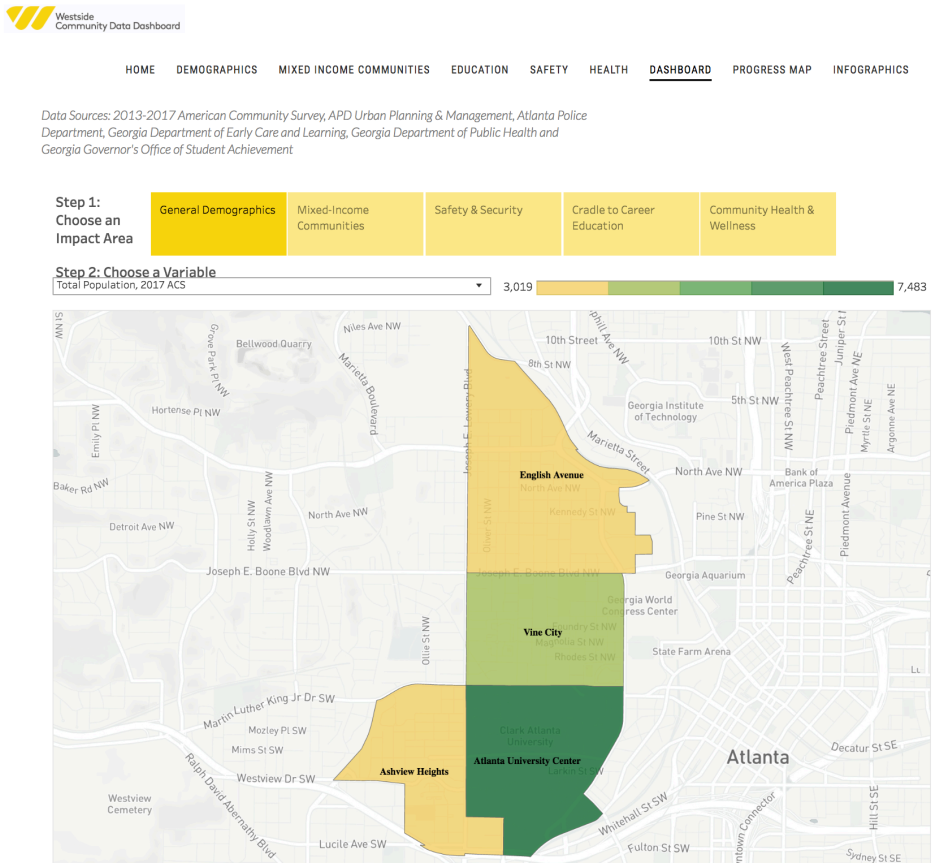
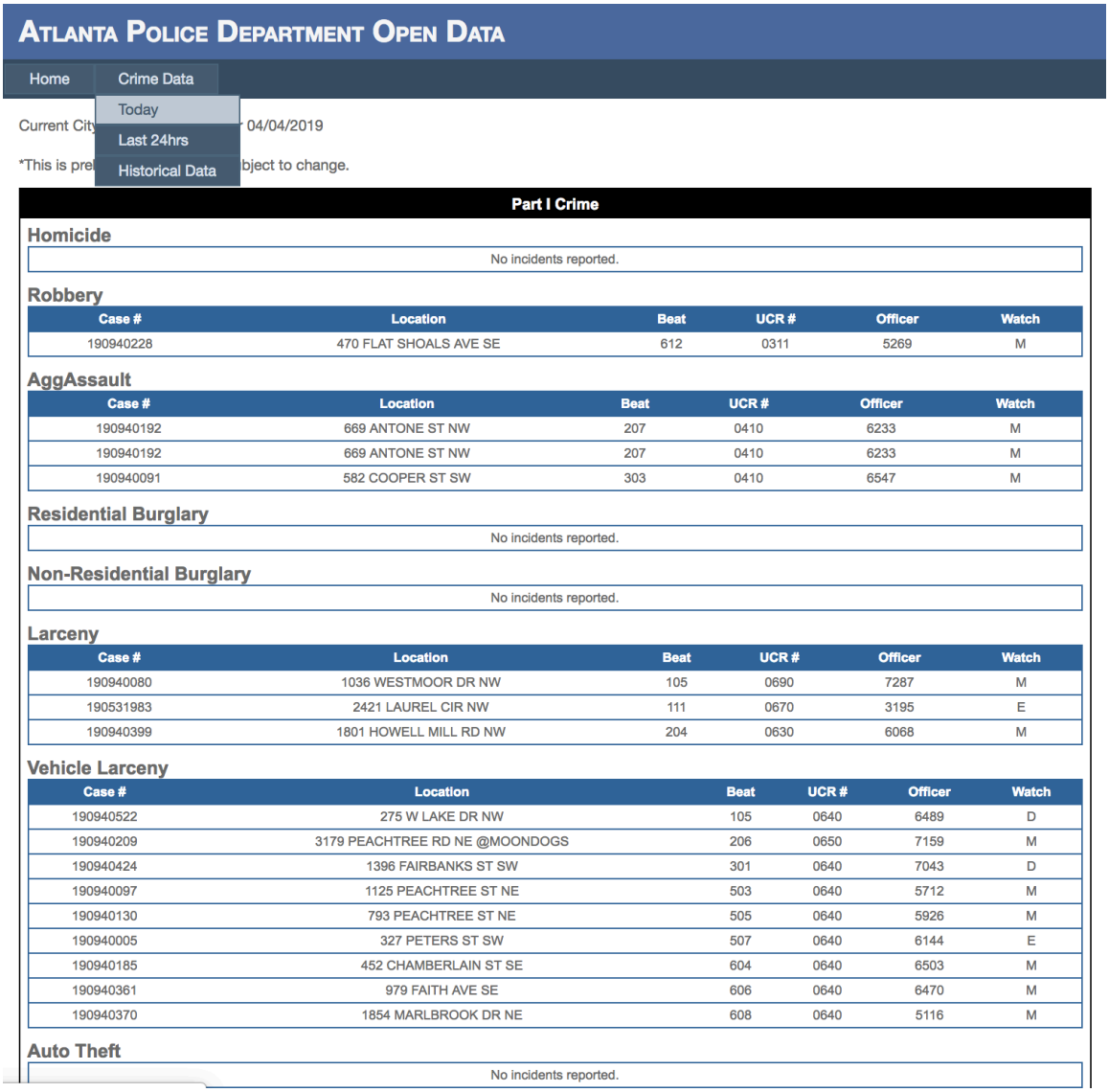


Figure 4: The Westside Future Fund Data Dashboard.

[8] https://www.dublindashboard.ie/pages/index



ATLANTA POLICE DEPARTMENT OPEN DATA

Home | Crime Data | Today | Last 24hrs | Historical Data

Current City ... 04/04/2019

*This is pre... bject to change.

Part I Crime

Homicide

No incidents reported.

Robbery

Case #	Location	Beat	UCR #	Officer	Watch
190940228	470 FLAT SHOALS AVE SE	612	0311	5269	M

AggAssault

Case #	Location	Beat	UCR #	Officer	Watch
190940192	669 ANTONE ST NW	207	0410	6233	M
190940192	669 ANTONE ST NW	207	0410	6233	M
190940091	582 COOPER ST SW	303	0410	6547	M

Residential Burglary

No incidents reported.

Non-Residential Burglary

No incidents reported.

Larceny

Case #	Location	Beat	UCR #	Officer	Watch
190940080	1036 WESTMOOR DR NW	105	0690	7287	M
190531983	2421 LAUREL CIR NW	111	0670	3195	E
190940399	1801 HOWELL MILL RD NW	204	0630	6068	M

Vehicle Larceny

Case #	Location	Beat	UCR #	Officer	Watch
190940522	275 W LAKE DR NW	105	0640	6489	D
190940209	3179 PEACHTREE RD NE @MOONDOGS	206	0650	7159	M
190940424	1396 FAIRBANKS ST SW	301	0640	7043	D
190940097	1125 PEACHTREE ST NE	503	0640	5712	M
190940130	793 PEACHTREE ST NE	505	0640	5926	M
190940005	327 PETERS ST SW	507	0640	6144	E
190940185	452 CHAMBERLAIN ST SE	604	0640	6503	M
190940361	979 FAITH AVE SE	606	0640	6470	M
190940370	1854 MARLBROOK DR NE	608	0640	5116	M

Auto Theft

No incidents reported.

Figure 5: Real time data from the Atlanta Police Department.

Several cities, counties, non-profit organizations and universities also build indicator dashboards that are geared more towards raising awareness and generating dialog about issues such as equity, environmental pollution, health, education and the like, which are of interest to specific communities. Within Atlanta, independent agencies like Neighborhood Nexus[9] (Figure 6) and the Atlanta Regional Commission[10] (Figure 7)

[9] https://neighborhoodnexus.org/

specialize in offering data dashboards and consultancy services for a variety of businesses, counties, and cities across Georgia. Outputs they create include dashboards, visualizations, reports, plans and predictions that are based on data that they collect. The Community Indicators Consortium[11] based out of Issaquah, Washington, is an organization that "advances and supports the development, availability and effective use of community indicators" by organizing educational opportunities, webinars, resources and annual summits that help communities and practitioners advance the practice and effective use of community indicators. Such events underscore the importance of understanding indicators and the manner in which they impact us and the cities we live in.

[10] https://opendata.atlantaregional.com

[11] https://communityindicators.net/

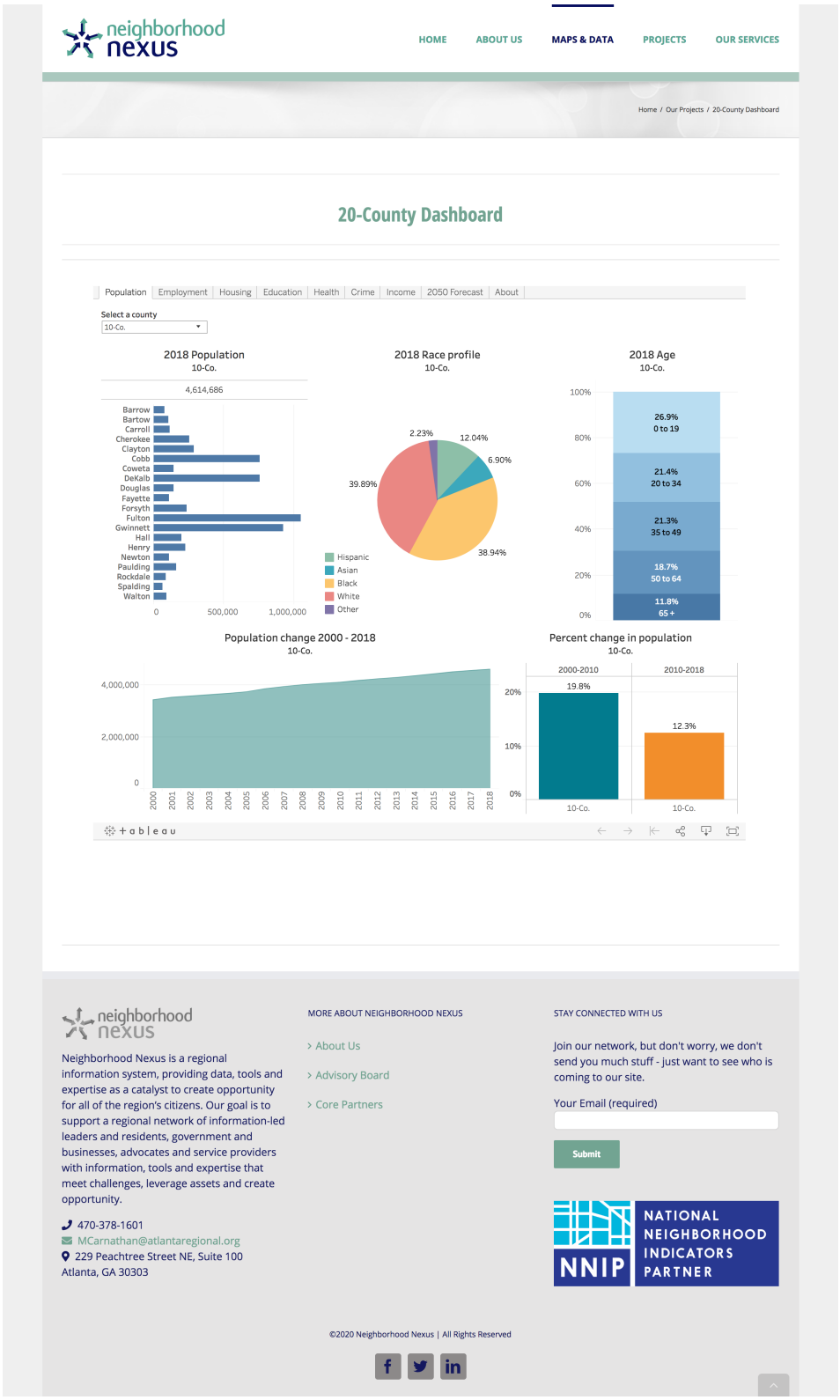


Figure 6: The 20 County Data Dashboard developed by Neighborhood Nexus.

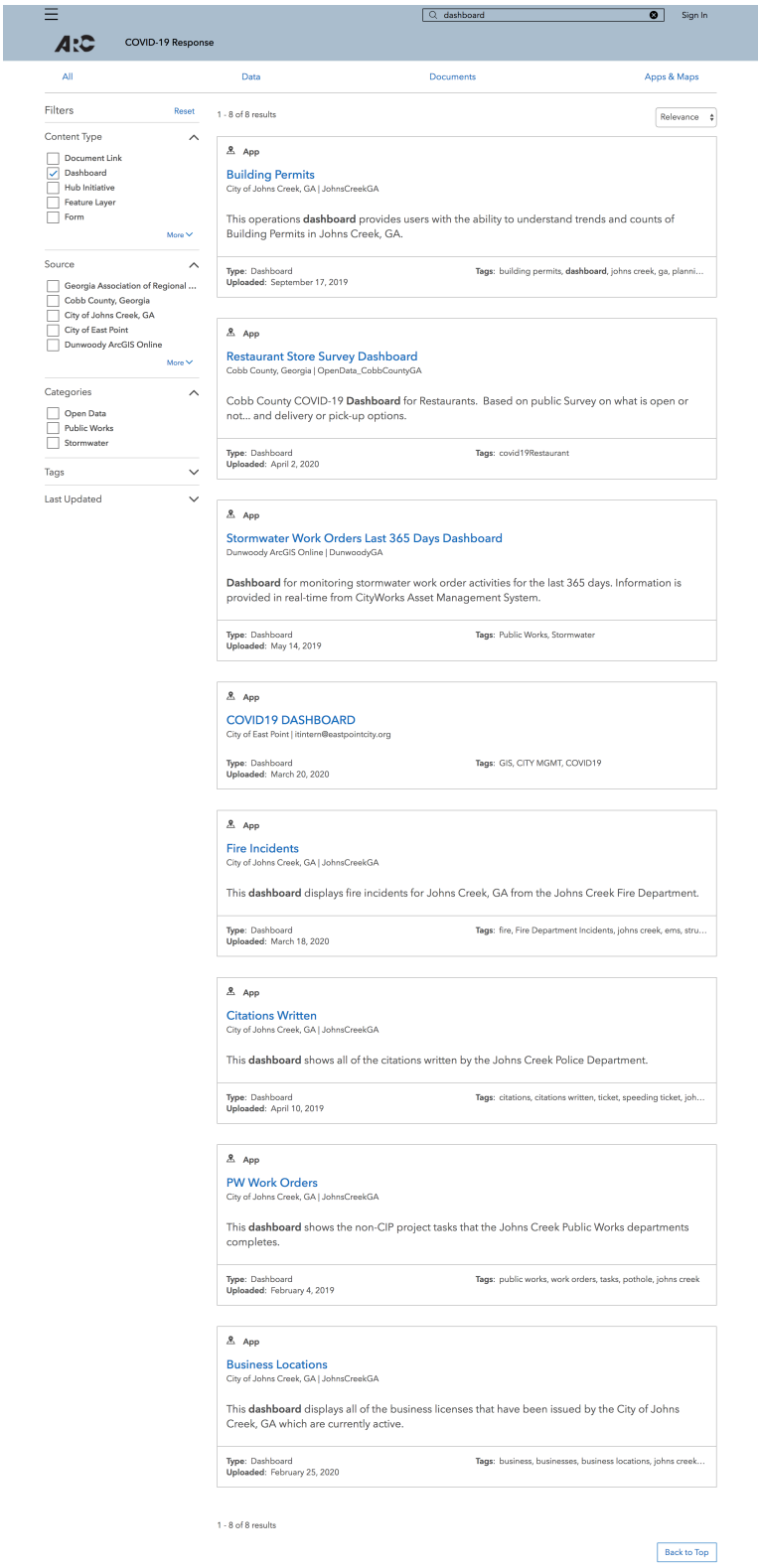


Figure 7: The multiple dashboards available on Atlanta Regional Commission's homepage.

Community indicator data dashboards are built with the explicit goal of transparency and fostering civic engagement. They are usually free and publicly accessible visual manifestations of data that are of common interest to the community within which they are situated. These dashboards typically consist of a variety of graphs, tables and maps that offer narratives about the community's demographics, education, public safety, health, jobs, quality of life etc. The emphasis is in making the process of moving from data to insight as intuitive and straightforward as possible, in a way that they are accessible to novice as well as expert users. Indicator dashboards in these instances are being treated as tools for community empowerment, as a way to allow concerned citizens and organizations to advocate for change within their communities.

But these indicators are not without their shortcomings. The aura of objectivity that these numeric representations are loaded with hide the messiness of the cities and communities they represent. As a technological manifestation of the complex and controversial elements that make up a city, indicator dashboards mirror issues of voice, power and representation that need to be investigated further if we are to understand their impact on communities (Kitchin, Lauriault, et al., 2015). Right from when the idea of an indicator dashboard is conceived, to when it is in the hands of users who are actively engaged in interpreting it, indicators are mired in a series of socio, technical and political negotiations and power struggles that need to be critically reflected upon if we are to truly understand how they work and the opportunities they enable (Mattern, 2014). Kitchin & Lauriault (2014) argue for a conceptual reimagining of data dashboards as "data assemblages – complex, politically infused, socio-technical systems that rather than reflecting cities, actively frame and produce them".

The complexity inherent in these dashboards is decipherable to only those who have the training, skills, knowledge and tools to make sense of the data, which are not evenly distributed within marginalized communities. By ignoring to make data equitably accessible to such marginalized communities, indicator dashboards could be seen as another in a series of injustices, which while being well intentioned, are actually perpetuating the data divide rather than bridging it.

The WCA Data Dashboard is one such indicator dashboard that brings together data from multiple sources under a common interface. The dashboard was built for residents of the Historic Westside neighborhoods in Atlanta and puts community data into context by providing an interface that allows users to cross-reference and visualize the data in multiple ways. By aiming for intuitive use and pleasurable experience, the dashboard highlights the objectivity of indicators, while simultaneously underemphasizing their underlying messiness, biases and shortcomings. The assumption here was that building such a dashboard would give communities in the Westside the data they need when advocating for change. *It is this assumption that I challenge in my dissertation*. Making data accessible is not the same as making it usable or equitable (Gurstein, 2003), and the argument holds in the case of data dashboards as well. In challenging the above assumption, my dissertation unpacks the infrastructural elements that are involved in building, using and maintaining equitable data dashboards for resource-constrained communities. I offer an analysis of how individuals in these communities are using the WCA Data Dashboard and whether their needs are being met. I reveal the barriers that individuals in these communities face when using the WCA dashboard and what we can do to ensure that the data infrastructures we build with and for them are aligned with their

values of data equity. I ground my investigation in the argument made by Mattern, Kitchin and Lauriault, that dashboards are data assemblages that need to be critically studied from a sociotechnical perspective in order to understand the human, material and organizational arrangements that make them work and opportunities they enable (Kitchin & Lauriault, 2014; Mattern, 2015).

The sociotechnical approach to unpacking the dashboard asks that we perform an infrastructural inversion (Bowker, 1994), where infrastructural elements that are often invisible and in the background are foregrounded and made the focus of analysis. This process of infrastructural inversion allowed me to observe the conflicted context within which the dashboard was born, and the influence this context continued to have on different elements of the dashboard (and vice versa).

The idea of a dashboard for equitable access to data for the Westside communities was initiated not from public workshops or deliberations, as is common within the fields of Community Informatics or Participatory Design, but from an idea that the Dean of the Ivan Allan College had based on her interactions with the community and the experience of using similar dashboards in her previous job. But the dashboard she used was made to manage student data in a university environment, which is quite different from the context in which the WCA dashboard was developed. While the WCA dashboard was well intentioned and initiated as one that would support social justice efforts by bringing about equity in how data are accessed and used, the question of what counts as just has been the subject of scholarly debate (Costanza-Chock, 2020; D'Ignazio & Klein, 2020; L. Taylor, 2017). My choice in analyzing this dashboard through a social justice and participatory design lens is so I can identify specific ways in which the dashboard might

live up to one of these perspectives of just. In doing so, I wrestle with the unjust and non-participatory origins of the dashboard, and in the process, help envision what a just and participatory approach to making data equitable for the Westside community would look like.

In adopting such a participatory and justice-based approach to analyzing the WCA dashboard, the research question I am asking is

> *What are the local values and infrastructural arrangements that are required to build, use and maintain equitable data infrastructures that enable communities to benefit from the publicizing of data through dashboards?*

To answer this question, I first performed an infrastructural inversion of the WCA dashboard. This led me to observe how the different sociotechnical elements of the dashboard came together, along with their inherent praxis and politics. The manner in which the Dean artfully integrated her resources to build the dashboard pointed to *infrastructural bricolage* as one way through which such dashboards could be put together within resource constrained environments. In addition to offering an ethnographic account of the infrastructural bricolage through which the dashboard came together and methodological specificity that enables the infrastructural inversion of data dashboards, my dissertation helps position data dashboards as an object of analysis within the larger field of information infrastructures. In observing the infrastructural breakdowns and specific ways in which the WCA dashboard failed to achieve its socially good goal of data equity, I argue that we need to move beyond designing for social good and towards designing for social justice. Such a move to designing for social justice would involve developing a local conceptualization of what it means to do justice to data.

I use the term *local* to refer to values of justice that are grounded in the experience of the community leaders one is designing for and with, rather than global values of justice that have been suggested in the scholarship around values in design, which I discuss further in the next chapter. This pivot to local values allows me to get to local formulations of data justice that might lead to human flourishing (Friedman et al., 2009; L. Taylor, 2017). I organized data literacy workshops in order to elicit these local values and also identify specific ways in which the dashboard was or was not meeting the goals of the community leaders. These workshops also pointed to other gaps and injustices that were inherent not only in the dashboard, but also in the facilitation of the workshops. This was a turning point in my dissertation, as it reminded me of how data and workshop activities that are meant to empower communities can in fact make them feel inadequate and unwilling to engage further. The data and activities that I asked the community leaders to engage with were a frustrating reminder of the injustices that the Westside community had experienced over the years. The local values of justice that I identified helped me counter some of these injustices in the next step of my research.

Countering the dashboard involved reimagining what it would take to ensure that its data could benefit the Westside communities, who did not always have access to the knowledge, time, skills and resources needed to make effective use of the data. In doing so, I rethought how the community leaders could access and make sense of data and the software platforms through which they would do so. This involved identifying data ambassadors who could guide other individuals and organization in the community to use these software platforms and also help them find, analyze and publicize data and stories around relevant matters of concern. Countering the dashboard in this way brought

attention to the *infrastructural arrangements*, or the human, material and organizational arrangements that are key to *building, using and maintaining* civic data within communities.

Being a student within a public research university came with its own restrictions and privileges, which shaped the kinds of access I had to individuals within the university and the wider community. As a researcher interested in partnering with the Westside community, I had to contend with the suspicion many community leaders had, which was a consequence of them being "underserved and over studied"[12] by students and researchers from neighboring universities who treated the Westside neighborhoods as a research lab. Much of my interaction with the community leaders was facilitated by faculty and staff at the Westside Communities Alliance (WCA), who were well recognized in the community. This dissertation is a consequence of the time, expertise and guidance I received from the WCA as well as the community leaders I interacted with.

Overall, answering this question helped draw attention to the human elements of civic data infrastructures, which led me to the title of my dissertation and to conclude that *all data are human*, and the way we do justice to them is by engaging with the human elements of their civic data. This argument is grounded in my experience of engaging with the different actors and networks that embodied the Westside community's civic data infrastructure and limited by my own positionality within this university-community context. My dissertation draws attention to the knowledges, skills, resources and infrastructural arrangements that communities can use to benefit from the publicizing of

[12] Chairperson, Neighborhood Planning Unit T (NPU-T), at a panel titled "Atlanta: Whose Data Is It Anyway?: Empowerment & Ownership of Community Research" at the Integrated Network for Social Sustainability (INSS) 2017, Atlanta, GA.

data through dashboards, which is just one aspect of the ongoing data revolution. Such an understanding will allow researchers and practitioners to move beyond platitudes about data leading to empowerment and instead focus on structuring environments and infrastructures that enable communities to equitably participate in the production and use of public data.

CHAPTER 2. RELATED WORK

The field of Human Computer Interaction (HCI) offers a wealth of scholarship to draw from when studying such data dashboards from a sociotechnical perspective. HCI is, generally speaking, concerned with how humans interact with computing technologies. Over the decades, the field has evolved to incorporate scholarship from an array of disciplines like Media Studies, Science and Technology Studies, Design Studies and the like, which has made the term HCI itself rather ambiguous. I situate my work within this evolving and diverse field of HCI by focusing on specific subfields that have concerned themselves with a critical analysis of the environments and infrastructures within which technologies can be developed for marginalized populations. I introduce some of these subfields in the next section and engage more deeply with them in each of the subsequent chapters.

The field of HCI started off with a concern for building models and theories of how humans were interacting with computers on an individual basis to now thinking about the broader impact such technologies have on society. This concern for the broader impact has led scholars to think about topics such as sustainability, justice, and care in the context of human computer interaction, while also incorporating feminist practices that consider the perspectives of collectives that are traditionally marginalized within such conversations. This has led to borrowing concepts from cognate fields like Science and Technology Studies, which considers more than just the actors who directly interact with technology and also consider the environments and infrastructures within which such interactions take place. A turn towards the infrastructural elements of technology

interactions is also the turn this dissertation takes, as it starts off with a focus on building the data dashboard and moves towards organizing its underlying human infrastructure. I review some critical scholarship that has guided my dissertation next.

2.1 Human Computer Interaction

Human Computer Interaction (HCI) emerged in the 1970's as a subfield of computer science that relied on cognitive science and human factors to design the early personal computers that were rising in popularity (J. Carroll, 2013). The field of cognitive science incorporated cognitive psychology, artificial intelligence, linguistics, cognitive anthropology, and the philosophy of mind to develop scientific models of how humans interacted with computers. Building on its industrial design roots, the field of human factors was first applied to study how humans operated computer systems in domains like aviation and manufacturing. As personal computers became the primary tools through which work got done in offices and factories, HCI shifted its focus from "human factors to human actors" (L. J. Bannon, 1986). This second wave of HCI relied on situated action, distributed cognition and activity theory to study the different contexts within which work got done (Susanne Bødker, 2006). Participatory design methods, prototyping and contextual inquiries were developed to focus more on the humans that did the work in these settings. The personalization and miniaturization of computer technologies, which led them into our homes also marked a shift in HCI that ushered in its third wave. Researchers in the third wave many of the second wave methods to examine the culture, emotion, values, and aesthetics of interacting with and designing wearable and tangible interfaces, augmented, virtual and mixed reality applications, and other computing technologies that have pervaded our social and professional lives. More recently, cities

and governments have also realized the importance of using computing technologies to become more efficient, transparent, resilient, so they can better engage with their citizens.

HCI uses various umbrella terms such as Digital Civics, PolitiCHI, HCI for Civic Engagement/Development, Community Informatics, Information and Communication Technologies for Development (ICT4D), Civic Tech among others to categorize research that seeks to broaden citizen participation in the design and use of technologies. While each of these terms bring focus to different aspects of intersection between civics and technology, collectively they advocate for the need to move out of the lab and into the real world to study and build technologies oriented towards strengthening democracies.

My work aligns most closely with research that falls under the umbrella of Civic Tech, Digital Civics and Community Informatics. Within each of these fields, I've been guided by scholarship that takes a critical and justice-based approach to studying and designing technologies for marginalized communities. I clarify what I mean by offering some definitions and examples of scholarship that has shaped my research in the next section.

2.1.1 Critical and Justice Based Approaches to the Design of Community Technologies

The field of Community Informatics (CI) is concerned with empowering communities to access and use Information and Communication Technologies (ICT's) in the service of community objectives (Gurstein, 2007). Gurstein goes on to say that

> *The objective of CI is to use ICT to enable the achievement of community objectives including overcoming "digital divides" both within and between communities. But CI also goes beyond discussions of the "Digital Divide" to examine how and under what conditions ICT access can be made usable and useful to the range of excluded populations and communities and particularly to*

support local economic development, social justice, and political empowerment using the Internet.

John M. Carroll & Rosson (1996) ran some of the earliest long-term studies around how communities in Blacksburg, Virginia engage with information through ICT's. Principles, frameworks and theories that came out of their work have guided more recent scholarship that falls under the umbrella of Civic Tech and Digital Civics.

Civic Tech takes a broad interpretation of technologies that "support both formal and informal aspects of government and public services" (Boehner & DiSalvo, 2016). These can include for instance, technologies required to bring clean drinking water to the low income residents of Detroit, Michigan (Sabourin, 2016; Schrock, 2018). Digital Civics on the other hand, is concerned with using digital technologies to empower citizens (Olivier & Wright, 2015). It "weaves together advances in data-driven service delivery, efforts to broaden participation in local governance and local public institutions through interactive systems, and thereof support deliberation and equity in local, national, and global civic enterprises" (Vasilis Vlachokyriakos et al., 2016). Researchers in both of these domains have considered how citizens are empowered through data (Asad et al., 2017; Vasillis Vlachokyriakos et al., 2015) and have also evaluated technologies that can be used to democratize data and make citizens part of the conversation (Aoki et al., 2009; Dickinson et al., 2018; Hara et al., 2013; C. A. Le Dantec et al., 2015; Maskell et al., 2018). Such work acknowledges the friction and polyvocality involved in using data for civic participation. Rather than treating citizens as mere consumers who are involved in transacting with government services, they focus on strengthening the relationship between government and their citizens. This involves considering not only the individual

technologies and systems through which governments and their citizens engage, but also the sociotechnical infrastructures upon which they are built (Korn & Voida, 2015). This concern for the infrastructural elements that underlie our digital civic information and communication technologies is the thread that runs through my dissertation as well as the projects I describe below. I outline some key projects that have informed my dissertation here and refer to them again as I discuss my results and contributions through the rest of this dissertation.

Balestrini et al. (2017) have proposed a six-part framework that focuses on the why and how of using technology to engage with local matters of concern. The six phases of the framework involve (i) identifying common matters of concern, (ii) framing the issue from different perspectives and identifying viable solutions, (iii) designing tools, interactions, management protocols and learning structures that can tackle the issue, (iv) deploying the technology so it can be tested, iterated and improved in situ, (v) orchestrating a sustainable engagement model that involve organizing events that help scale the technology to wider audiences, and finally (vi) reflecting on how the resulting outcome helped the community achieve its goals and sharing best practices that might benefit future engagement efforts. The authors stress the centrality of humans being involved through all the phases as well as the role of the researcher, which is to "explain, fire-fight and help, but not to control or manage". This framework and the conceptualization of the role of the researcher was helpful as I structured my own engagement with the Westside community leaders.

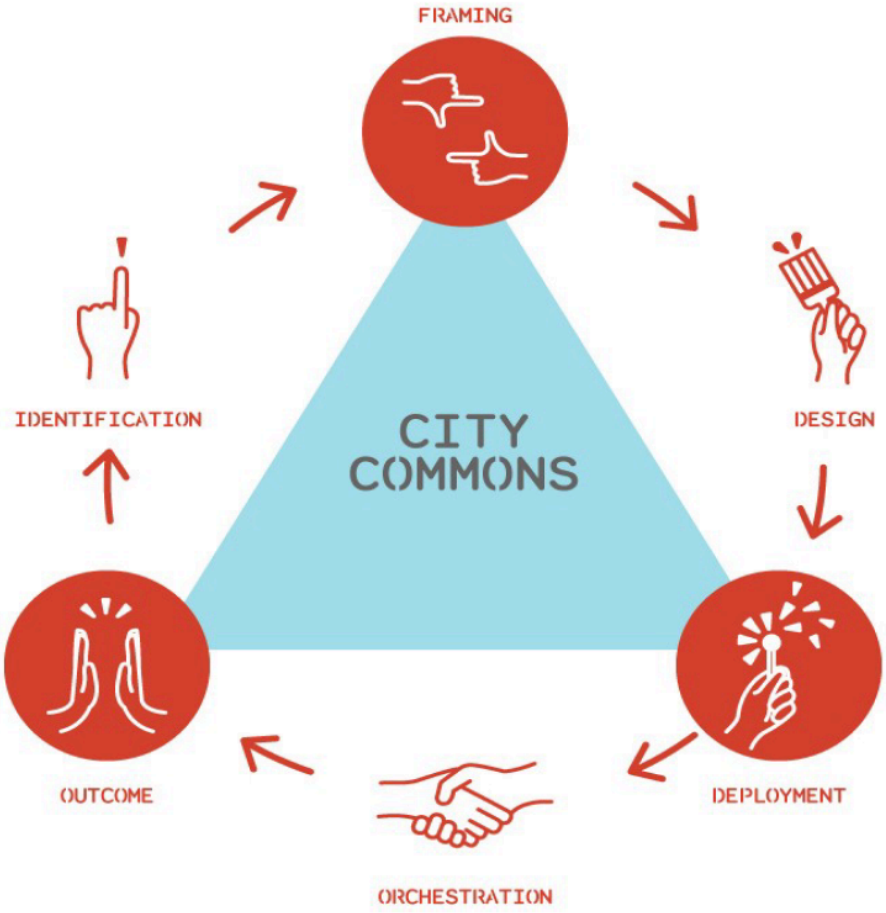


Figure 8: A city-commons framework for citizen sensing by Balestrini et al. (2017)

Over the decades that J. Carroll et al. (2018) have been working with communities, one of the most common barriers they experienced was a lack of data literacy "The challenge of participatory design in contemporary community informatics is chiefly one of creating a self-directed and sustainable process of continuous learning" (J. Carroll & Rosson, 2007). They advocate for overcoming the data literacy barrier through formal and informal programs as necessary. This requires reimagining the social as well as the technological infrastructures within which communities 'access', 'interpret' and 'use' data (J. Carroll, Shih, et al., 2015; Gurstein, 2011).

In acknowledging the need to work with local communities to co-design tools and processes that would enable data novices to access, interpret and use data, Puussaar et al. (2018) developed the Data: In Place (https://app.data-in.place/) platform. Recognizing that "Making data available is not the same as making data usable for everyone" the authors built scaffolds into the platform that allowed citizens to perform data science without the help of expert data analysts. Users start off the interaction by drawing boundaries around the place they are interested in. Once they have done this, the platform fetches all the datasets it can access about that geographic area. Fetching the latest data every time the user draws a geographic boundary allows the app to always display the latest datasets for that region. The user then has the option to select one of the datasets, upload their own or even request a new dataset to be added to the platform. Once the dataset is selected, the user is offered options to customize the kind of visualization they would like to produce. The tool uses specialized vocabulary and data analysis procedures that would only be accessible to those with experience in data science and visualization. Such tools can be made accessible to individuals without such experience through data literacy initiatives as described by (J. Carroll et al., 2018), which was something I considered when developing the WCA dashboard as well.

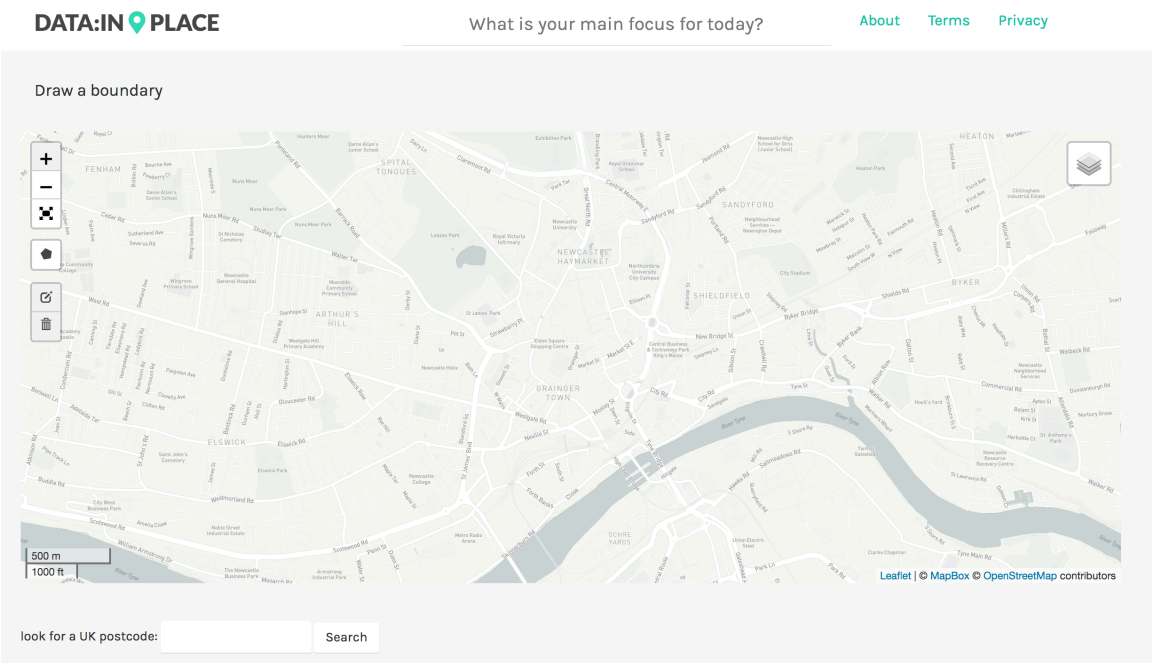


Figure 9: Data: In Place app by Puusaar et al.

The Data: In Place app was developed by researchers at the Open Lab at Newcastle Upon Tyne using a host of JavaScript mapping and visualization libraries like MapBox, Leaflet, Raw Graphs and server-side logic in NodeJS, PostgreSQL, and PostGIS. This implies that the team is also responsible for maintaining the app going forward to ensure that nothing breaks when any of the libraries are updated or servers become unavailable. What happens when these researchers leave the project and how it might be handed over to the community is not discussed.

Some scholars use the phrase ‘Research in the Wild’ to refer to HCI research that has moved out of the safe and controlled lab environment and into the streets and homes of communities where technologies are being developed and evaluated in situ (J. Carroll & Rosson, 2013; Chamberlain et al., 2012; Rogers, 2011; Rogers & Marshall, 2017; N. Taylor et al., 2013). One of the concerns driving their scholarship is about strategies that can help ensure that community projects can continue to thrive once researchers have left

the field. Taylor et al. (2013) point to three issues when designers hand over technologies to communities. First, the technologies that the designers build are usually prototypes rather than finished products. They are not designed with long term use and support in mind. Second, using these technologies require the designer to have created the capacity within the community to do so. This includes creating content, interest, or service models that will allow the community to continue to use the technology once the designer is not involved. Third, creating and sustaining technology and use require human and financial resources that change over time. Each of these were issues I encountered during the development of the WCA dashboard and have more to say about them in Chapter 4.

The newfound availability of previously inaccessible data combined with cheap computing power and storage devices have fuelled the hype around Big Data and Open Data systems. This hype has prompted scholars to also turn their attention to overcoming the 'data divide' that is impacting marginalized communities use data (Andrejevic, 2014; Boyd & Crawford, 2012). Scholars devote much of their attention to the equitable distribution and consumption of data and information technologies in order to benefit communities on the wrong side of the data divide. Such approaches to overcoming the data divide are grounded in notions of democracy, social good, empowerment, equitable distribution of power.

Community data refers to "data pertaining to a community and its locale, that is, gathered, analysed, interpreted and used by members of a local community" (J. Carroll et al., 2018, p. 2). They argue that involving citizens in the analysis of community data has the potential to transform citizens from being mere data points to data analysts who are capable of contributing to the local democratic processes. In order to do this, we need to

go beyond focusing on access to data and think about the barriers that prevent citizens from effectively using the data (Gurstein, 2011). Gurstein (2011) further argues that we should strategize to go beyond concepts like 'digital divide' or 'data divide' to concepts of 'effective use' by arguing that (Gurstein, 2003, 2011)

> *"what is significant is having access and then with that access having the knowledge, skills, and supportive organizational and social structures to make effective use of that access and that e-technology to enable social and community objectives"*

Effective use for him is "the capacity and opportunity to successfully integrate ICTs into the accomplishment of self or collaboratively identified goals" (Gurstein, 2003). He differentiates between the process of interpreting or understanding open data and making effective use of it. J. Johnson (2014) argues that a focus on access tends to downplay how data often carry assumptions of privilege and ignores the fact that data users might have limited technical capabilities.

Critical data scholars have documented how open data initiatives have ended up empowering the already empowered rather than the marginalized. One example comes from Benjamin et al. (2007) who document how the ``effort to digitize land records in Bangalore allowed the middle and upper class to usurp land from the marginalized poor by taking advantage of the gaps and irregularities in their ownership documents. Gurstein (2003, 2011) and O'Reilly (2011) have warned that if precautions aren't taken, open data can easily be appropriated as a tool to be used against the marginalized (O'Reilly, 2011). Visions of smart cities that are built on concepts like Open and Big Data have ignited the

debate around the implications they have for issues of trust, accountability, governance, citizenship, power, privacy and inclusion (Dencik et al., 2019).

Such issues form a subset of concerns that Feminist scholars in HCI have taken up as they propose new methods, theories and approaches that centre the voices of women and other marginalized communities when designing data technologies. Such scholarship is grounded in feminist standpoint theory (Harding, 2004), which argues that all knowledge production and experiences are partial and subject to acts of power in which the voices of the marginalized are suppressed. Such acts of power are compounded and lead to a form of structural oppression, often experienced by individuals who exist at the intersections of marginalized races, genders, and classes (Crenshaw, 2018). When combined with religion, nationality, location, a lack of literacy, generational poverty and the prison pipeline, this leads to what Collins (2000) calls the matrix of domination, which are interlocking systems of oppression that shape peoples lives (Collins, 2000).

Feminist scholars have proposed a number of principles, guidelines and frameworks that can guide intervention within such systems of oppression. I introduce some foundational ones here and engage with them and others more deeply in Chapters 5 & 6.

Bardzell (2010) has argued for a feminist approach to HCI that takes pluralism, participation, advocacy, ecology, embodiment, and self-disclosure as qualities to start with (Bardzell, 2010a). Dombrowski et al. (2016) argue for a social justice orientation to design practice that is rooted in a *commitment to conflict, a commitment to reflexivity and a commitment to ethics and politics* (Dombrowski et al., 2016). To help design practitioners avoid designing for the status quo that often reproduce systems of oppression and social inequities, they propose that we design for designing for

transformation, recognition, reciprocity, enablement, distribution, and accountability. A commitment to these strategies, qualities and principles are also reflected in the Design Justice Principles, which is a living document that lists a series of ten principles that helps rethink "design processes, centers people who are normally marginalized by design, and uses collaborative, creative practices to address the deepest challenges our communities face" (Costanza-Chock, 2018, p. 2). Data Feminism, a concept put forth by D'Ignazio and Klien (2020), seeks to understand how feminist scholarship can inform an approach to data science and ethics (D'Ignazio & Klein, 2020) that is grounded in issues of social justice. Their approach to not taking data as given and the collapsing long established binaries have spurred reflective and critical representations of data. I take up their feminist principles of data visualization (D'Ignazio & Klein, 2016) further in Chapter 6.

In taking such a justice based approach to data, Taylor (2017) argues that the data revolution has so far only been a technological one and asks technologists to think harder about the social justice implications their technologies have on individuals and communities (L. Taylor, 2017). Since such technologies unequally impact those who are already marginalized, it is crucial to consider how intersectional qualities like "race, ethnicity, religion, gender, location, nationality, socio-economic status – determine how individuals become administrative and legal subjects through their data and, consequently, how those data can be used to act upon them by policymakers, commercial firms and both in combination" (L. Taylor, 2017, p. 3). A rights-based approach to data justice is based on individual and collective rights like the right to free speech and communication, the right to privacy and the right to data protection. However, Taylor (2017) argues, that such a framing is problematic because it assumes that any violations

would be easy to spot and can be redressed at an individual level. But violations to data-based rights aren't always clearly visible and easy to respond to because they tend to take place across groups rather than individually.

Taylor contends that there are at least three main approaches to data justice. One, put forth by Johnson (2014) argues for "subsuming the question of open data within a larger question of information justice" (J. Johnson, 2014, p. 1). Despite the best intentions of all the stakeholders involved in open data projects, the process of formatting, recording, making retrievable and relatable, and communicating that information is inherently value laden, biased and structured to empower those who are already empowered (Gurstein, 2011). "Injustice in, injustice out" is the phrase Johnson (2014) uses to describe a variety of open data projects that try to live up to the ideals of distributive justice and accountability promised by such efforts. Any efforts to remedy the injustices perpetuated by open data projects will fail unless one has access to, and the ability to influence the underlying data, structures, databases, and processes and politics that govern them.

Second, Heeks and Renken (2018) suggest different ways in which data justice can be conceptualized in the development literature (Heeks & Renken, 2018). They build on UN's Universal Declaration of Human Rights to suggest procedural, instrumental, and distributive as the mainstream approaches to data justice in the field of human development. In order to judge if data are procedurally just, one needs to know if data were collected with the appropriate consent, if data are processed in a consistent manner that produce accurate results that are free from bias and if any errors produced can be corrected. The instrumental view on data justice is concerned with the outcomes that result from using the data. From this perspective, the specifics of obtaining consent and

processing data in a fair manner are subordinate to the impact that the data produce. Data is said to be instrumentally just if it abides by what is morally, legally and culturally right at a specific time and place. Distributive justice asks that data be distributed in a fair and just manner. When seen from a rights based perspective, it would imply that everyone should have equal right to data access, privacy, ownership and representation (J. A. Johnson, 2016).

In critiquing these mainstream approaches Heeks and Renken (2018) also point out that these approaches tend to assume data as 'Big' that are controlled and acted upon by governments and large corporations, while ignoring the 'small' data that individuals/communities need and value. These mainstream approaches also ignore the sustainability of data as a justice issue, which Heeks and Renken believe need to be considered more explicitly. They ask that we look beyond theoretical principles of justice and consider what people are capable of doing and what they actually do with the data. This means that have having access to data in insufficient for data justice. Invoking Amartya Sen's capability approach (which I explain later), they ask that we also need to consider the capabilities (data and technological literacy, agency, power etc.) and functionings (what they achieve through their capabilities) of the people involved. Human agency and motivation are also crucial drivers of what data justice means to these people. Such capabilities, functionings, motivations and agencies do not develop in a vacuum, but from how society is structured.

A third perspective comes from Dencik et al. (2016) who, in the wake of the Unites States' and other countries' shift to surveillance capitalism, consider the conditions under which data should not be distributed (Dencik et al., 2016). Calling for greater integration

between technology and social justice activists, Dencik et al. (2016) suggest a data justice framework that treats surveillance as a core political issue that everyone in our society needs to be concerned with rather than just a small set of experts. This involves paying greater attention to those in power who determine how our digital infrastructure is organized along with thinking about their "security, autonomy, dignity, fairness and sustainability" (Dencik et al., 2016, p. 9).

Given that these three perspectives have different visions of what is considered just or what should and should not be done when it comes to data, Taylor (2017) offers a framework based on human needs that can help reconcile these different perspectives. She brings these three perspectives together through the three pillars of her data justice framework: *visibility, engagement with technology* and *non-discrimination*.

Visibility refers to those who are routinely ignored in the data, while also acknowledging their desire for privacy. It is well known that the census is not an accurate representation of the number of undocumented immigrants in the US. These immigrants do not respond to the census surveys from fear of becoming visible to the state, which could lead to deportation. *Engagement with technology* asks that those who are represented in the data also have the right to share in its economic benefits, while giving them autonomy to choose if and how they are included in such technologies and data markets. Taylor cites the case of Google trying to use India's biometric population database, Aadhaar to identify its next billion users (Aulakh et al., 2016) as a case of data injustice being done to those represented in the Aadhaar database. This is because the potential economic benefits of this data are reaped by Google only and the Indian population have no say in determining the terms for how, why and with whom their data

gets shared. An approach based in data justice would give the Indian population the freedom to control their engagement with specific technologies that make them politically or commercially visible. The third pillar, that of *non-discrimination* asks for power to identify and challenge any bias in data use and the freedom to not be discriminated against. A case in point would be the one from Bangalore, India mentioned earlier, where the land of the poor was usurped by the middle and upper class thanks to the many irregularities in the ownership documents that were modified to favor those with money and influence (Benjamin et al., 2007).

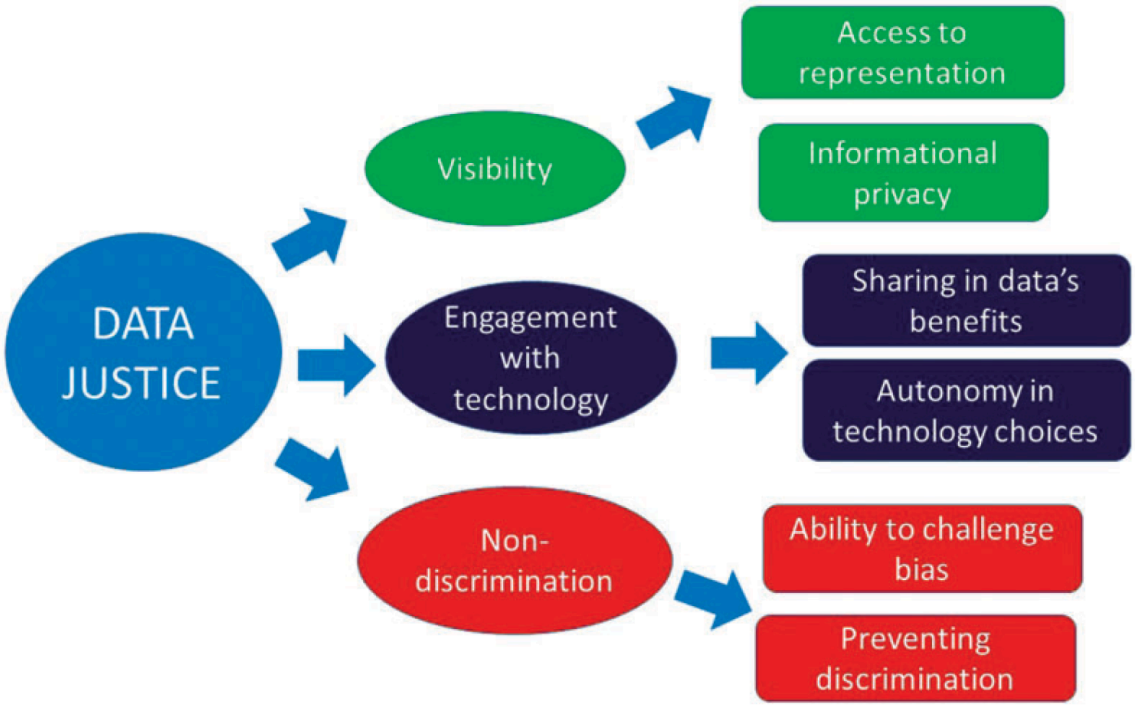


Figure 10: Three pillars of data justice, from Taylor (2017)

Taylor takes this framework a step further and offers a way for it to be operationalized. She uses Sen's Capability Approach as a way to operationalize her framework to guide future research and debate in the field.

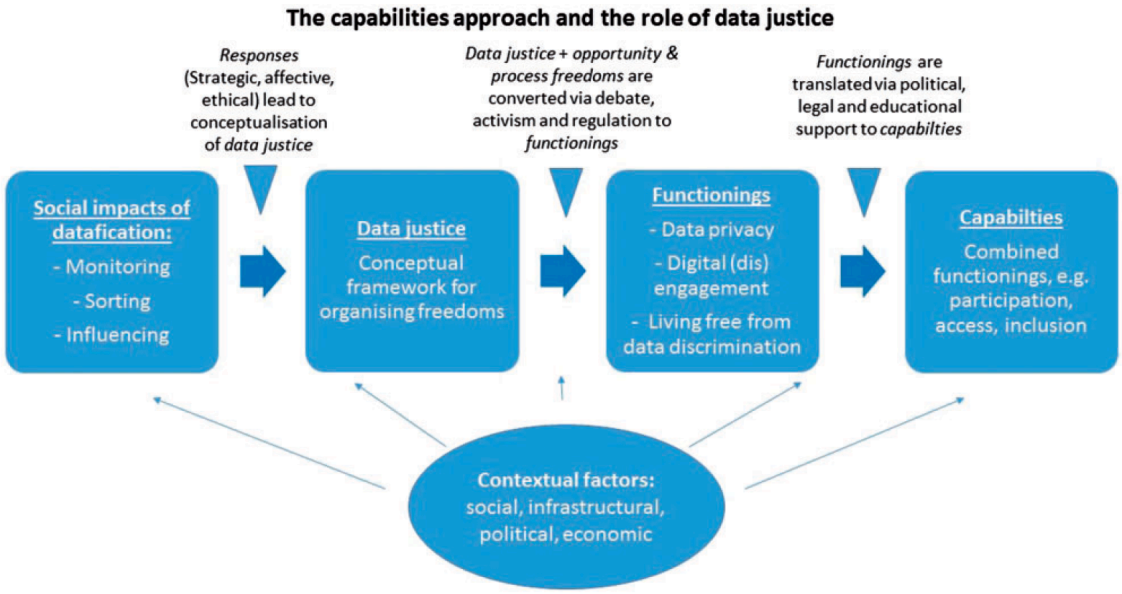


Figure 11: A capabilities approach to data justice, from Taylor (2017)

When assessing individual well-being or social arrangements, policies or proposals about social change in society, the capability approach asks that we focus on two key idea, functionings and capabilities. Functionings are 'beings and doings', that is, different states of human beings and the activities they can undertake. Examples of 'beings' would be being educated, being illiterate, being part of a community or being happy. Examples of 'doings' would be taking a test, taking part in a strike, buying groceries or driving a car. Capabilities are a person's real freedoms or opportunities to achieve functionings. Thus, while being educated is a functioning, the opportunity to be educated is its corresponding capability. The capability of humans to achieve these functionings also depends on conversion factors such as their personal (intellectual capacity, physical condition, education, skills etc.), social (social norms, public policies, gender norms etc.) and environmental (climate, physical location and infrastructure etc.) characteristics. This approach accounts for human diversity and is able to consider the experiences of the

marginalized rather than those of the average person (L. Taylor, 2017). A capabilities approach asks that we take a practice-based approach to data justice, to consider individual capabilities that can lead to valuable functionings in society. Using such a capabilities approach in conjunction with the framework Taylor has suggested allows us to arrive at a conceptualization of data justice that that can go beyond the utilitarian conceptualization of benefits, to one that is based on values that can lead to the flourishing of the humans involved. Taylor argues that such a 'local' conceptualization of data justice that is contextually dependent is what we should be striving for, rather than one that is global and based on the normative view social justice. Inspired by this idea of the local, my dissertation is an effort to generate one such local conceptualization of the values that describe what data justice means for a group of marginalized communities in Atlanta. I draw on the values in design scholarship in HCI to elicit these values.

2.1.2 The Turn to Values in HCI

The field of HCI, which during the 1970's was mostly concerned with the needs of the computing systems has evolved to make values more central to the conversation on how technological systems must be designed. This pivot to values, they argued, was necessary as it allowed us to design technologies that would enable human flourishing rather than just meeting their needs. While needs refer to something that is required and maybe even important, like food, water, or shelter, values refer to that which is more fundamental or substantial to the individual or group, like care, justice, humor, accountability etc.

Value Sensitive Design (VSD) is an approach that calls for a "theoretically grounded approach to the design of technology that accounts for human values in a principled and comprehensive manner throughout the design process" (Friedman et al., 2009, p. 349).

The authors prescribe a three-part framework to investigate the conceptual, empirical and technical design of computational systems. As part of this three-step process, practitioners in the VSD tradition are guided by set of twelve ethics based human values: Human Welfare, Ownership and Property, Privacy, Freedom from Bias, Universal Usability, Trust, Autonomy, Informed Consent, Accountability, Identity, Calmness, and Environmental Sustainability.

These universal values, although effective at getting practitioners to consider the ethical issues involved in the design of computational systems, obfuscate more culturally specific interpretations of local values. Borning & Muller (2012) have therefore encouraged researchers to be more explicit in letting their research participants speak for themselves, rather than taking on the role of interpreters or reporters of these contextual situations. In doing so, they ask the researcher be more reflexive of their own values and biases, including the influence they have in giving voice to these values of these participants.

Similar critiques of VSD comes from Le Dantec et al., (2009), who, in questioning these values argue that such a discursive list does not allow for the discovery of more local values that are specific to the context of design. They instead ask for more prescription for the methods that can be used to elicit these values from local contexts rather than prescriptions as to what these values actually are.

Parvin et al, publishing as JafariNaimi et al., (2015) argue that design needs to go beyond the identify/apply logic of values, wherein values are accurately identified through relevant research methods and then applied to the design of technologies that embody those values. Values, they argue, serve as hypotheses within problematic

situations, which develop through inquiry and cannot be treated as pre-established formulae. “Rather, values serve as *hypothesis* by which to examine what the situation is, what the possible courses of action are, and how they might transform the situation. Judgements, including design judgements, are the outcome of practical, intellectual, and emotional interaction with situation that are indeterminate or puzzling. Values are not *applied* to situation; rather, values *serve* situations as hypotheses” [emphasis in original] (Parvin et al, publishing as JafariNaimi et al., 2015, p. 97).

But these values dynamic and need to be negotiated throughout the design process (Grönvall et al., 2016). Building on design values that ground the field of PD (L. Bannon & Ehn, 2013; Kensing & Greenbaum, 2013), Grönvall et al. (2016) use the notion of infrastructuring to negotiate these changing values and account for the unbounded nature of the settings within which community based PD takes place.

My process of discovering local values started off by engaging the Westside community through data literacy workshops, which I describe in Chapter 5. These values then served as hypotheses during the redesign of the Westside’s civic data infrastructure, which was my was of countering the injustices inherent in the WCA dashboard. This process of infrastructuring, which I describe in Chapter 6, reveals how values are situated and need to be negotiated throughout the design process.

My quest to discover local values of data justice was further guided by STS and feminist perspectives on technologies in which scholars emphasize how knowledge is contextual and situated in their places of origin (Dourish, 2004; Haraway, 1988; Loukissas, 2016; L. Suchman, 2002; Vertesi & Dourish, 2011). In claiming that All Data Are Local, Loukissas (2019) brings this idea of localness to bear on different collections

of data like those from the Arnold Arboretum, the New York Public Library, NewsScape and Zillow. His claim that "data are cultural artifacts created by people, and their dutiful machines, at a time, in a place, and with the instruments at hand for audiences that are conditioned to receive them" (Loukissas, 2019, p. 2) has been helpful in unpacking the sociotechnical black boxes of the Westside's data infrastructure as well. Indeed, the title of my dissertation can be thought of as a portmanteau of his book *All Data Are Local* and the tech firm *All Tech is Human*, which is working to incorporate multiple stakeholders and perspectives to produce more inclusive, multidisciplinary and participatory technological change in society. I engage with other concepts from Loukissas like *data settings* and the *Civic Data Guide* in Chapter 7.

This chapter offered an overview of specific literature within the field of HCI in which this dissertation is situated. Inspired by scholarship that specifically takes a critical and justice-based approach to designing data and information systems for communities, this dissertation offers a local and situated description of what it takes to design such systems for the Westside communities in Atlanta. The methods I used to do this are primarily ethnographic and borrow from the Scandinavian tradition of participatory design and the sociotechnical study of information infrastructures. I describe these methods in further detail in the next chapter.

CHAPTER 3. METHODS

The field of Participatory Design (PD) has a rich tradition of such situated analysis and intervention, which is what made it an excellent source for methodological inspiration for my dissertation. In an almost parallel epistemological development, the field of Information Infrastructures (II) also put forth a set of recommendations for the study of large-scale information infrastructures. There is much that is common between these methodological traditions and I use this chapter to situate my approach within them.

I study the dashboard not as a tool or an interface that needs to be evaluated, but as a sociotechnical assemblage which is influenced by and has an influence on the context within which it is being developed. The dashboard is part of an already existing information infrastructure which includes the datasets, websites, institutions, policies and organizations that are required to bring it to fruition. The scholars in the field of II's were pioneers in taking an ethnographic approach to study the influence that these elements have on the technology and vice versa (Bowker, 1994; Bowker & Star, 1999; S. Star & Ruhleder, 1996). In specifying how we ought to conduct an ethnography of infrastructures, they draw on classical anthropological concepts of constructing, extending and bounding the field of inquiry by engaging with the different elements of the infrastructure, pursuing the phenomenon, following connections and discovering discontinuities. Bowker and Star (1999) also introduced the concept of infrastructural inversion, which involves foregrounding the often-invisible elements and processes underlying information infrastructures. I combined these approaches with PD methods

that are specifically concerned with issues of power and democracy when designing technologies with communities. I rely on the theoretical frameworks from the PD tradition in order to make sense of my field studies and design collaborations.

I situate my methods by offering a brief overview of how the practice of ethnography evolved. Ethnographic practice within HCI has a number of flavors and influences, and the one's I focus on are those within the traditions of PD and II's. I start off with an overview of the field of PD and its relationship with ethnography. After an explanation of these relationships, I then go on to describe my own methods and the manner in which they aligned with these traditions.

3.1 Participatory Design

As described in the previous chapter, the field of HCI has sought to consider not just the usability needs of its end users, but also their social, political, environmental, civic, democratic and emancipatory needs. Just as research in HCI was moving towards novel ways of responding to such needs, researchers in Scandinavia were examining the idea of workplace democracy, which came about largely with the introduction of computers into the work environment. Driven by union workers who wished to be involved in the decision making process that determined how computers would change the future of their work environment, Scandinavian researchers put forward the idea of Participatory Design (Kensing & Greenbaum, 2013).

Early Participatory Design (PD) researchers were critical of the social, political and economic values that had so far guided the design and development of workplace technologies in a top down manner, which completely ignored the needs of the workers. These researchers argued that the workers who would be using the technologies should

have a say in how the technologies are designed to support their work. In doing so, they established a blueprint for projects to come by employing a mutual learning approach that treated users and designers as experts in their own fields who could benefit from each other's domain knowledge (Nygaard & Terje Bergo, 1975).

In the first generation of PD projects, Scandinavian scholars worked with union workers in factories in order to tailor workplace technologies to their needs. There were three broad categories of projects that came about through this participatory approach to design. 'Knowledge Strategy' projects were committed to bridging the knowledge gap that existed between union workers and the computer system developers. The Norwegian NJMF[13] project from 1970 to 1973 (Nygaard & Terje Bergo, 1975) was one of the first such projects that inspired many others within this space. What started off as a fairly traditional workplace design project, in which the researchers set the agenda and sought feedback from the workers through interviews, ended up being flipped based on feedback received from the trade unions. Researchers were concerned that the knowledge they produce would be of no practical relevance to the workers to whom it mattered the most, and therefore decided to include them in design negotiations from the very start. This process resulted in a set of teaching materials and actions that allowed workers to learn new knowledge that would enhance the way they worked with computers on the factory floor.

'Design and Intervention' projects, the second category, went a step further than just sharing knowledge between the different actors, to actually getting workers involved in the design of systems alongside the developers. The UTOPIA project was commissioned

[13] NJMF is the Norwegian acronym for Norwegian Iron and Metal Workers Union.

to develop technological alternatives for graphical workers in the Swedish and Danish newspaper industry (S Bødker et al., 1987). The researchers setup a laboratory where workers and developers would work together to prototype and test out new tools and technologies for digital printing. Florence was another design and intervention project that aimed to build computer systems for nurse's daily work at two hospitals in Oslo, Norway (Bjerknes & Bratteteig, 1987).

Anthropologists at the Xerox Palo Alto Research Center (PARC) in Palo Alto, California, carried out what are considered a third category of projects that influenced the Participatory Design movement. Lucy Suchman, Janet Blomberg , Eleanor Wynn among others at PARC were interested in the ethnographic analysis of relations between work and technology (Blomberg et al., 1993; L. Suchman, 1987; L. A. Suchman, 1983; L. Suchman & Wynn, 1984; Wynn, 1979). The theories and cognitive models they produced ushered a second generation of PD projects, which were focused on the design and development of collaborative workplace technologies.

The third generation of PD projects focused on large scale systems and tried to bring in heterogenous stakeholders in an inter-organizational context (Pilemalm, 2018; Pilemalm & Timpka, 2008). The guiding philosophy here was that large-scale workplace technologies impact a variety of stakeholders who need to be considered in the design of such systems. Such efforts had to contend with issues like contacting these different stakeholders, discerning their conflicting needs, balancing such extensive participation with the demanding timelines of these projects (Grudin, 1993).

More recent efforts in PD have seen researchers move out of the workplace and into communities that are looking to engage with their governments through ICT's

(Information and Communication Technologies). The concept of the end user has also been the subject of scholarly reinterpretation and has changed as PD moved out of the workplace and into communities. Seeing the value that the perspectives of the end users can bring to a project, researchers have tried to engage them throughout the PD process rather than just in the ideation phase. Concepts like design after design and participatory infrastructuring were developed (which I will explore further in section 3.5) in order to draw attention to the role of end users and designers in sustaining PD projects in the wild (Ehn, 2008; Karasti, 2014).

Community based PD "foregrounds the social constructs and relations of groups in settings that include, but go well beyond, the formal organisational structures commonly foregrounded in more traditional workplace studies" (DiSalvo et al., 2012). Communities can be formed based on geographic proximity, shared identities around race, gender, ethnicity, or even shared interests, hobbies and practices. When working with these communities, a designer needs "to be able to reflect upon not only activities in the design process, but also upon the multiple intentions and interpretations that build the analytic lens of the research or design project" (Mörtberg et al., 2010, p. 107). This call for reflexivity in ethnographic practice has been taken up by scholars working in the HCI and PD tradition (Dourish, 2006; C. Le Dantec & Fox, 2015; Rode, 2011) and have influenced the manner in which I report my findings here. I discuss this notion of ethnography within the HCI and PD tradition that I am drawing from next.

3.2 Ethnography

Ethnography has been a central component of PD since its inception. Ethnography has its roots in anthropology, where it was first used to study the everyday lives of non-Western

societies (Agar, 1996). The practice of ethnography is founded on offering a descriptive understanding of phenomena in their everyday settings and taking a holistic view that prioritizes the perspectives of its members (Blomberg et al., 1993). It has traditionally been concerned with observing the social life, cultural institutions, religious customs, political practices and other aspects of everyday community life in order to produce meaning.

Ortner, (2006) describes ethnography as "the attempt to understand another life world using the self – or as much of it as possible – as the instrument of knowing". Dourish, (2014) further unpacks this definition by drawing attention to three key ideas – the focus on the holistic life of the participants suggested by the terms *life worlds*, the focus on *the self as the instrument of knowing* and the partial nature of the enterprise as suggested by the phrase *as much of it as possible*. What these authors suggest is that, in an attempt to understand and make meaning of the world of the other, the ethnographer is limited by their own physical, intellectual and cultural sensibilities. This understanding of ethnography has brought greater attention to notions of subjectivity and reflexivity on part of the ethnographer, a recognition of the unboundedness of the sites and phenomena under investigation along with the realization that the best any ethnographer can do is to offer their interpretation of the evidence that they gather. Each of these ideas have generated the critique and controversies that have shaped the practice of ethnography over the decades and their uptake in the field of HCI and PD.

When adopted by researchers in design-oriented fields like HCI, PD and CSCW, the agenda was to use ethnography as a tool to understand the everyday lived experience of the people they were designing tools for. These fields are more concerned with using

ethnography as a way to intervene in the situation and affect change, than with producing descriptive accounts. Dourish (2014, 2006) asks that we move away from this narrow conceptualization of ethnography that is empirical and *only* produces "implications for design", to one that is also conceptual and "may work best not by providing answers but by raising questions, challenging perceived understandings, giving silenced perspectives voice, and creating new conceptual understandings" of sites and situations that might help us see them in a new light. Scholars have different visions of how ethnography ought to serve design, which has led to several interpretations of ethnography in design.

In the first wave, called *Ethnography for Design (E4D)*, ethnography was used to generate a set of requirements that could be used to design relevant products and services for users. Work within the first wave of HCI research described in the previous chapter generally borrows from the tradition of E4D. The second wave focused on using *Ethnography to study Design (E2D)*, which accounts for the reflexive turn in design where the focus in on studying how designers go about their work. Scholars in the field of Participatory Design have contributed a lot to this wave of E2D. The third wave of ethnography, called Design Ethnography by some and Design Anthropology by others, blurs the boundaries between design and ethnography so the designer/ethnographer is "actively immersed and engaged in a setting where people are either designing artifacts, producing artifacts, or introducing artifacts into a social and cultural context". My ethnographic approach similarly tries to unpack the different artifacts of the Westside community's data infrastructure by participating with the community in different contexts and capacities. Such an approach has previously been used within Participatory Design

and the study of Information Infrastructures. I offer a brief survey of these traditions before describing my approach in more detail.

3.2.1 Ethnography in the Participatory Design Tradition

The core commitments of ethnography outlined above are well aligned with the goals and aims of Participatory Design (Blomberg & Karasti, 2013). The focus of PD projects was and continues to be on advocating for those who are socially, politically and economically disenfranchised to be included in designing the technologies they use. Such projects employed a mutual learning approach that treated users and designers as experts in their own fields who could benefit from each other's domain knowledge (Nygaard & Terje Bergo, 1975). This need for knowledge exchange required researchers to introduce accessible spaces where designers, workers and other users could come together to voice their concerns and exchange ideas. This required the researchers to come up with new methods, tools and techniques that would allow workers to collaborate with designers to collectively produce visions of how the technology would change the work environment. The goal of such two-way participation was always towards emancipatory ends for the disenfranchised workers (Ehn, 2017), where concepts like democracy, situation-based actions, mutual learning and equalizing of power relations were of critical importance. As the field matured, democratic values that were focused on developing alternate visions of technologies through situated actions, mutual learning and equalizing of power relations between the different stakeholders were drafted into a set of guiding principles for the field of PD (Kensing & Greenbaum, 2013).

For (Blomberg & Karasti, 2013), ethnography and PD have been explored through three influential research programs, which they label as *Ethnography and Participatory*

Design in reflexive relation, Ethnography as a component of Participatory Design methodology, and *Ethnography to inform design*. Research carried out by the Work Practice and Technology group at Xerox PARC, where they critically analyzed existing work practices in order to design case-based prototypes which were again tested in actual work environments. This back and forth between the anthropologists and the computer scientists, between making sense of existing workplace practices and making something new is an example of *Ethnography and PD in reflexive action*. As ethnography became more commonplace within PD, it was systematically integrated into the collaborative design of new technologies with diverse stakeholders. The MUST methodology proposed by (Kensing, Simonsen, & Bodker, 1998) is one example of *Ethnography as a component of PD*. The MUST method asks that designers themselves be involved in interviewing and observing users in the field rather than relying on the ethnographer for design inputs. While the work of intervention, which is the focus of design, and that of description, which is what ethnographers are concerned with, are fundamentally different, "combining the two approaches and iterating between them has been an effective way in learning about the organisation and has been an important resource in generating realistic visions of future use" (Kensing, Simonsen, & Bødker, 1998). Simultaneously, researchers at Lancaster University carried out a series of ethnographic studies with the goal of informing the design of information systems for air traffic (Harper & Hughes., 1993; J. A. Hughes et al., 1992), law enforcement (Shapiro et al., 1991) and financial services (Blythin et al., 1997; Rouncejield et al., 1994). Preferring the separation of ethnographers and designers, these studies proposed 'debriefing meetings' and 'presentation frameworks' as ways to communicate findings between designers and ethnographers.

This tradition of using *Ethnography to inform design* has resulted in several variations like concurrent ethnography, quick and dirty ethnography, evaluative ethnography, re-examination of previous ethnographic studies (J. Hughes et al., 1994) and rapid ethnography (Millen, 2000).

Such ethnographic explorations have also led to debates about positioning design within ethnography and vice versa, as well as the separation between the two (Crabtree, 1998; Crabtree et al., 2009; Dourish, 2006). Foregoing such debates, Halse. et al., (2010) propose the Design Anthropological Innovation Model in which they argue that we "abandon the idea that the field of use is a place to visit and to be known, and that the design studio is a privileged place for invention" so we can "unleash a greater potential of combining anthropology and design". There is a productive tension in the coming together of the terms design and anthropology, which give it a distinctive focus on concrete practice and reflective action. Where design is future oriented and focused on intervention and change, the field of anthropology "systematically investigates the past to understand the present", with the focus being on cultural interpretation and production of theory (Otto & Smith, 2013, p. 4). This tradition of design anthropology is grounded in PD and takes it further by [emphasis in original] "emphasising the theoretical or cultural *frameworks* and the socio-political *contexts* within which both field studies and design collaborations are conducted and understood within the design process approach" (Smith & Kjærsgaard, 2015, p. 75). It is this tradition of Design Anthropology that I borrow from and develop further in the rest of this chapter.

As technology development moves from the design of customized software towards the integration of platforms and infrastructures, contemporary PD has also taken a longer-

term perspective that includes 'design after design', where end users' appropriation of platforms and infrastructures in use is considered a continuation of design. Such an orientation towards long-term infrastructuring has also been taken up by ethnographers of large-scale information infrastructures. I discuss this parallel tradition of ethnography and its relevance to my dissertation in the next section.

3.2.2 Ethnography of Large-Scale Information Infrastructures

Initially advocated for by Star (1999), the ethnography of infrastructures calls for a mixed methods approach to studying infrastructure systems as a network of relationships. In pointing out the salient features of infrastructures, Bowker & Star (1999) emphasize how infrastructures are embedded within their surrounding sociotechnical arrangements and are transparent to use as they do not have to be invented each time across multiple events or sites. Such infrastructures are learned as part of a membership, in that they require outsiders to become familiar with its objects, which are built on an installed base using established standards to plug into other infrastructures and only become visible upon breakdown. Collectively, these infrastructures are shaped by and shape conventions of practice and politics within specific communities (Bowker & Star, 1999; Star & Bowker, 2002).

The relational quality suggests that infrastructures mean different things to different people and revealing how they work involves studying "boring" aspects like standards, lists and specifications that are in the background but critical to the everyday working of infrastructures. Edwards et al. (2013, p. 5) emphasize that infrastructures have a "modular, multi-layered, rough-cut character… [that] are not systems, in the sense of fully coherent, deliberately engineered, end-to-end processes. Rather, infrastructures …

consist of numerous systems, each with unique origins and goals, which are made to interoperate by means of standards, socket layers, social practices, norms, and individual behaviors that smooth out the connections among them". Bowker (1994) refers to the foregrounding of these background elements of work as an "infrastructural inversion", which usually involves some combination of design and anthropological research methods including but not limited to interviews, observations, workshops, historical/literary/systems analysis, usability studies etc.

Karasti & Blomberg (2018) extend the work of Star and Bowker by focusing on additional qualities of information infrastructures; their emerging and accreting nature as well as intervention and intentionality that address the role design plays in infrastructuring. Karasti (2014) coined the term *infrastructuring* as a way to draw attention to the processual, ongoing and temporal manner in which such information infrastructures unfold. Infrastructuring refers to the "ongoing and continuing processes of creating and enacting information infrastructures" (Karasti & Blomberg, 2018, p. 234). Given that infrastructures accrete over time through specific actions taken by its actors, focus needs to be on the temporal aspects of information infrastructure or the manner in which they emerge over time. They suggest inverting infrastructures by investigating their moments of breakdown, following how their members engage in infrastructural inversion and by following the material and technical traces left behind by the different elements in the infrastructure.

Ethnographic approaches which include interviews, observations and document analysis are well suited to performing infrastructural inversions (Blomberg et al., 2003; Star, 1999). Such methods help elicit different points of view on the work involved in

bringing the infrastructure together. The ethnographic study of infrastructures must also reflexively examine the ethnographers role in constructing the field of inquiry. As Karasti & Blomberg (2018) argue, "The notion of constructing the field offers a nuanced approach to investigating information infrastructures, one that is inherently partial and one that aims to increase awareness of this partiality and the fieldworker's role in constructing the field".

An individual interaction with infrastructure can be a unit of analysis that can be connected to theory and used to shed light on different parts of the broader infrastructure. Connecting the dots between these individual encounters I examine the different human, material and organizational elements that help shape and sustain the Westside community's data infrastructure.

My encounters with the Westside community took place through interviews, focus groups, and workshops I organized as well as the public festivals, events and meetings that I attended. I took photos and wrote down notes through each of these encounters which helped offer more context when I was analyzing the data. I used Grounded Theory Method to make sense of the data I collected.

Grounded Theory Method (GTM) is explicitly concerned with the iterative interpretation of data to produce theory, which in turn informs the collection of additional data that can be used to create more robust abstractions and theories. Originally proposed by Glaser & Strauss (1967), the GTM has evolved over the years and take on different instantiations like Charmaz' (2006) constructivist methodology, Clarke's (2003) situational analysis and Locke's (2000) grounded theory approach in business and management some of which have been widely adopted by the HCI and CSCW

community (M. Muller, 2014; M. J. Muller & Kogan, 2010). Star and her early collaborators (Bowker & Star, 1999; S. Star & Griesemer, 1989; S. L. Star, 1985, 2012) have been some of the most prolific users of these methods and have produced a wide variety of concepts, theories and classifications that have guided future scholarship.

The Grounded Theory Method of analysis involves a stepwise iterative cycle of data collection and theoretical analysis. The process begins in data and involves (i) Collecting data and applying open, axial and selective coding practices to classify the data into categories, (ii) generating substantive theory based on the data and generated categories, (iii) writing memos that analyze the coding process and make the knowledge known, (iv) theoretical sampling, which is an abductive form of reasoning that is "strategic, specific, and systematic" (Charmaz, 2006) and (v) constant comparison "of data with data and of data with theory" (M. Muller, 2014) until all the categories are saturated and can be explained by corresponding theories (Charmaz, 2006; Corbin & Strauss, 2008; Dick, 2000; Glaser & Strauss, 1967; M. Muller, 2014; S. L. Star, 2012). During my interactions with the Westside community, as I went through similar stages of data collection and theorizing, which eventually led to my description of the human infrastructure of civic data. I describe my three-step process below and also explain how each step in my research informed subsequent steps.

3.3 Step 1: Sociotechnical Unboxing of the Data Dashboard

The call to unravel the messiness or the socio-technical complexity underlying numeric representations like dashboards have come from different scholarly traditions. Following the call by boyd & Crawford (2012) and Crawford et al. (2014) for a more critical and theoretically informed analysis of data, scholars have put forward a situated, reflexive

and contextually nuanced epistemology for Critical Data Studies (Dalton & Thatcher, 2014; Kitchin & Lauriault, 2014). In re-appropriating data dashboards as 'data assemblages' Kitchin & Lauriault (2014) urge us to consider the "the technological, political, social and economic apparatuses and elements that constitutes and frames the generation, circulation and deployment of data". Studies in metrology have taken an STS approach to expose the hidden social, economic, cultural and political implications of indicator systems within different fields like public health, statistics, climate change, justice and the like (Rottenburg et al., 2015). Researchers within CSCW have used ethnographic methods to examine the people and processes involved in creating large, multi-sited information systems like Enterprise Resource Planning software (Pollock & Williams, 2009; Williams & Pollock, 2012), scientific research infrastructure (Karasti et al., 2006; Ribes, 2014), hospital healthcare management systems (Hanseth & Monteiro, 1998), disease classification systems (Bowker, 1996) to generate theories on how information infrastructures can be conceptualized, built and maintained.

My ethnographic investigation of the WCA dashboard started in the summer of 2016, as a fellow in the Data Science for Social Good (DSSG) program at Georgia Tech (GT). The program is modeled after the DSSG program at the University of Chicago, where teams of students partner with local organizations who are working on socially relevant problems, to understand their data related needs and implement solutions. The fellowship brought together students from across the country, to spend 10-weeks in Atlanta working closely with pre-identified community partners.

My DSSG team chose to design and develop the Public Safety Module of the WCA dashboard. Being a three-person team, we did everything from data gathering, data

cleaning, visualizing, web development, while also attending neighborhood planning meetings and Public Safety Committee meetings to learn more about the community. During the 10-week fellowship, we attended two Neighborhood Planning Unit (NPU) meetings, and two public safety committee meetings. We presented the dashboard and the work we were doing to representatives from the police department and the community at one of the public safety committee meetings. We also presented our work along with the other DSSG teams at two public events that were organized by the university.

After the internship ended, I met with and conducted multiple semi-structured interviews with seven staff and faculty members at GT who were involved in the dashboard's early conceptualization. These interviewees in turn put me in touch with three community residents who were familiar with and also involved in planning the WCA dashboard. In total, I conducted ten interviews in this phase of the project that ranged from 30 to 60 minutes in duration. I also had access to previous reports and documents that the dashboard team had prepared as part of their community engagement and development efforts.

My status as a PhD student in a well-endowed public institution helped me approach community residents, but also brought with it, baggage from previous relationships between the community and the university. Managing these relationships and being sensitive to how I, as an outsider was perceived by the community was a key part of my ethnographic research. Part of doing this work before the work (C. Le Dantec & Fox, 2015) required me to build a familiarity with the community and their matters of concern. I did this by being present at community events, NPU meetings and the places that the dashboard was being featured at. This included dashboard launch events, public

presentations/demos, award ceremonies, conferences and the like. I also volunteered at the local boys and girls club and at local festivals that celebrated the neighborhood and its achievements. Following the actors through the community (Pollock & Williams, 2009) and the publics that formed through the various matters of concern (C. Le Dantec & DiSalvo, 2013) helped reveal the sociotechnical alignments of the community's data infrastructure.

Even though I was a student in the same college that was developing the dashboard, getting access and permission to do my research was not entirely straightforward. The Dean who initiated and funded the dashboard was always weary of outsiders usurping her work and taking credit for it. This was complicated by the relationship between the university and the Westside residents, which, as I explained previously, was not without incident. The Dean and the WCA were concerned that my project would be another in the list of projects that tried to study these neighborhoods and identify solutions to problems the community already knew about. I had to meet with the dashboard leadership team multiple times to present my project and convince them of the integrity in my research, its benefit to the community and the dashboard itself. The dashboard was going through its own set of management and funding issues, which further complicated matters. I first approached the dashboard team with my idea to perform an ethnography of the dashboard in the spring of 2016, but it wasn't until the fall of 2017 that we formally signed a Memorandum of Understanding (included in Appendix A) where both sides agreed on our respective roles, responsibilities and deliverables for the future of this work.

These power dynamics made some parts of the infrastructure easier to access than others and had direct implications for how our field site was constructed and studied. I

recognize these limitations and reflexively foreground them in my reporting of this project (Dourish, 2006; C. Le Dantec & Fox, 2015; Rode, 2011). My reflections in Chapter 4 are a result of an inductive analysis of the notes I wrote and the interviews I recorded during this engagement process, the core of which lasted for about a year between the summer of 2016 and spring of 2017. Although I did not formally interview the dashboard leadership team again, I continued to be involved in their team meetings and planning sessions. This allowed me to observe the ongoing changes with the dashboard and also clarify some of my previously collected data. I open coded my notes to create a set of themes and categories that stood out during the engagement process (M. Muller, 2014). I then used axial coding[14] to organize these codes into broader themes that helped explain the praxis and politics involved in building such dashboards. I compared these themes with existing theories describing infrastructural inversions of information infrastructures and data dashboards. This iterative process helped me recognize the different infrastructural elements, their power, influence, agencies and capabilities that allowed them to bring the dashboard together. This process also revealed data literacy as a critical barrier that needed to be overcome to ensure that the dashboard could serve the data advocacy needs of the Westside community. Thus, the next step in my research involved organizing data literacy workshops in the Westside.

3.4 Step 2: Organizing Data Literacy Workshops

I came across a range of data literacy skills and capacities within the cohort of community leaders I interacted with. As described further in Section 4.1, all community

[14] Axial coding is a qualitative research technique that involves relating data together in order to reveal codes, categories, and subcategories ground within participants' voices within one's collected data (Allen, 2017).

leaders I spoke with were experts in the needs of their communities. They knew about the issues their neighborhoods were grappling with and the data that was needed to draw attention to them, but not everyone knew where to find the data or how they could be analyzed, processed and presented as a coherent story. Individuals who were comfortable using data analysis and presentation tools like Tableau, D3, and Power BI were few. Considering these variety of data literacies, I proposed the data literacy workshops to the WCA as an effort to bridge the gap between the community leaders' knowledge of the social context and the tools that can be used to find data, analyze and present them in relevant formats. The WCA agreed to support my proposal considering that it built upon what they were hoping to do with their dashboard training sessions as well.

As captured in the figure below, residents of the Westside neighborhoods have been the subject of several disparaging media and research reports, which have made them reluctant to participate in more research studies. It was important that I built the relationship with the community leaders and establish a level of trust (C. Le Dantec & Fox, 2015), so they would be willing to share their time and expertise with me. This process of relationship building helped me contextualize the workshops to the specific matters of concern that were relevant to the Westside.

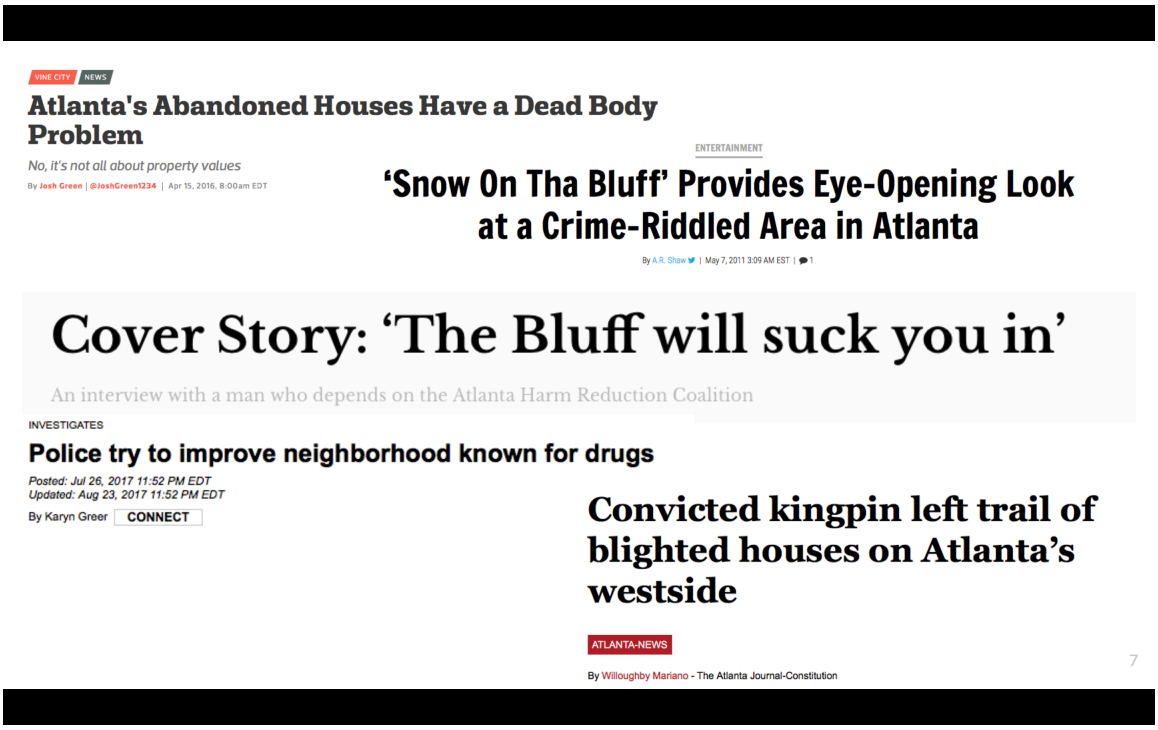


Figure 12: A selection of new stories about the Westside

I organized eleven data literacy workshops between March 2017 and July 2018. The first two workshops in March served an exploratory purpose, to understand the kinds of activities and datasets that would pique the community's interest. These two workshops were organized at the Westside Resource Center in partnership with the WCA. They also helped me recruit participants for these workshops by sending emails to non-profit organizations that had expressed interest in learning about and using community data in their daily operations. The six participants who attended these first two workshops came from non-profit organizations who used data about the community's education, demographics, income, jobs etc. to track and measure the performance of specific programs in the neighborhoods. They had a broad range of skills, with some being only moderately familiar with Microsoft Excel while others had also used professional software programs like Tableau. After the workshops, I interviewed three of the six

participants based on an interview guide that I developed and got approved by Georgia Tech's IRB (Institutional Review Board). The questions included general ones about the participants' background and skills with respect to data, as well as more specific ones about what they would like to learn at future workshops. Findings from these exploratory sessions and interviews helped me tailor the remaining workshops to the needs of the community. Aligning the workshop context with the context of use for the non-profits was my attempt at making the workshops more relevant to the participants.

While participants in these exploratory workshops were eager to learn about data, they also expressed the need for an incentive to participate. Kayser-Bril (2016) has similarly argued that it is hard to get people excited about data literacy if they do not have an incentive or a reason to learn it. The Dean agreed and wanted to make an offer of support to participants that was meaningful and that would reflect the significant time investment on their part. She also wanted this stipend to convey the university's commitment to the work of community organizations and resident advocates who were participating in the study. We therefore setup the remaining workshops as a set of four weekly sessions, where each organization could send up to two participants and would be compensated with a $400 stipend for every participant that completed at least three of the four workshops.

This stipend was significant and influenced the kind of participants who signed up. In the first round of workshops for example, there were four participants who admitted that using data was not part of their everyday job, but they were interested in picking up new data literacy skills. One participant even admitted that she was there so she could raise

money for her organization. Having expected this, I structured the workshop activities so they would appeal to participants with a broad range of skills and motivations.

3.4.1 Structure of Workshops

I structured the workshops as a series of four that would teach participants the basics of data storytelling with the CWK data dashboard. These workshops took place in February/March and June/July of 2018 at two easily accessible locations; a public library and a community food bank. I organized an additional make-up workshop at one of the participating organizations office for those who missed a session in the second round.

Figure 13: Workshop setup at the Atlanta Community Food Bank, Feb 2014.

3.4.2 Choosing activities

Rather than creating new data literacy workshop activities, many of the activities I chose were based on the worksheets and tools created as part of the Data Culture project (D'Ignazio & Bhargava, n.d.-a). The goal of the Data Culture project is to go beyond boring spreadsheet-based trainings and use an arts and crafts based data storytelling approach to build a data culture within organizations. The project was developed by working with a variety of non-profit organizations who were hoping to build a data culture within their organizations. This involves giving the entire organization, and not just the IT group or leadership, access, tools and literacies required to derive insight from the organizations data and tell stories with them (Bhargava, 2017). The authors were inspired by the critical and emancipatory potential of data literacy when building out these activities (D'Ignazio & Bhargava, 2016), which is what makes these activities a good fit for my research as well.

The Data Culture project website includes sample data, worksheets, and how to videos in which the authors describe how the activities can be used within workshop settings. The activities (all taken from DataBasic.io and included in Appendix B) I borrowed included Building a Data Sculpture, Asking Good Questions of Data, Sketching a Story with Word Counter, Try to Convince Me and Remix a Visualization using a sample dataset. The activities used craft materials, sketching and role-playing activities to build a basic understanding of what counts as data, the kinds of questions one can ask of them and how they can be used to sketch stories and make convincing arguments.

3.4.2.1 CWK dashboard activities

For the activities that involved the data dashboard, I used an information hunt activity that was put together by the dashboard leadership team (included in Appendix B). This

was initially developed as a way to evaluate if users would be able to find relevant data on the dashboard. The team used this activity in workshop settings where participants were asked to use the dashboard as a reference when answering the listed questions. The activity offered the team cues into how the data on the dashboard and its modules could be better organized to improve data discovery. The activity worksheet included questions corresponding to each of the modules, such as

Percentage of the population in Atlanta that is age 5 to 17 years?

APS school with the highest 7-12 dropout rate and one with the lowest

Year construction began on MARTA rail system?

I used this information hunt activity to generate discussion about the dashboard's usability, its data, and how they both could be improved to suit the work that the community leaders were doing. I did not set a strict time limit and encouraged participants to clarify their doubts and ask us any questions they had. I thus used the information hunt activity not only as a measure of usability or navigational efficiency, but also as a tool to scaffold critical discussions about the dashboard and its infrastructural black boxes.

For the final activity which asked participants to create data narratives with the CWK dashboard, I took inspiration from the Data 101: Data Visualization, Data Literacy and Storytelling workshop (Monk & Tutt, n.d.) format, which introduced participants to the different types of data stories (Bhargava, n.d.-a) (connection, comparison, change, personal and factoid), and asked them to create one of their choice. The focus here again was to use data available on the CWK dashboard to find and create a story that interested the participants. I chose the data storytelling approach as it closely matches the way non-

profits use data within their own work, either when advocating for change or seeking out grants (C. Clarke, 2009; Erete et al., 2016).

3.4.2.2 Participants

A total of seven participants from five different nonprofit organizations attended each of the first four workshops in February/March of 2018. I followed up these workshops with interviews of six participants who had agreed to share more details about their data needs and issues but did not have time to do so at the workshops. Each interview lasted for about one hour and was transcribed prior to analysis. I did not collect demographic data from these participants. My inferences about their data and computer literacy skills are based on the discussions we had during the interviews and the workshops. The participants were representing the following organizations

1. **SWEEAC (The Southwest Ecumenical Emergency Assistance Center)**, which is a collective of churches that offer emergency food, clothing, health, hygiene and other resources to needy families in the south west region of Atlanta. Stephanie and Rhonda who represented SWEEAC at the workshops had basic computer skills but did not use data as part of their jobs. They were interested in learning more about how they can use data to contribute to their organization's mission.
2. **TryCope (Try Reaching Youth Through a Community of Parenting Effectiveness)** is a community of neighbors in Vine City who offer academic support, extracurricular activities, and mentoring opportunities to empower their children to grow valuable connections among community members and excel intellectually, emotionally, socially, and professionally. Bluma attended

all the workshops and Carol dropped out after two. Both required minimal assistance when using the computer. It wasn't clear how much data they interacted with as part of their jobs.

3. **Sisters Action Team**, takes a holistic approach to promote personal wellness among residents of the Westside neighborhoods. Their organizational efforts include programming and networking opportunities that tend to the physical, emotional, spiritual, mental, economic, and environmental health of the community. Kimberly and Dayana who attended were involved in collecting data from neighborhood residents and using them to create presentations for a wide range of audiences (city officials, other non-profits, neighborhood residents etc).
4. **WAWA (West Atlanta Watershed Alliance)** represents the African American communities in Atlanta that are most effected by environmental stressors and works to improve their quality of life by protecting, preserving and restoring the community's natural resources. Isabel was the only participant from this organization had no experience with data and was there to learn about new tools that she might be able to use to access and make sense of the data her organization collects.
5. **Historic Westin Heights / Bankhead Neighborhood Association, Inc**, represents residents of the Historic Westin Heights and Bankhead neighborhoods. This organization does not have a dedicated web page, but their Facebook page gives the impression that their primary goal is to make sure that the voices and concerns of their residents are heard at the decision-

making circles. Stella had lots of experience collecting and making sense of data, but this was a long time ago. Owing to her age and deteriorating eyesight, she relied on other participants to read out what was being displayed on the computer screen.

An additional nine participants from six different nonprofit organizations attended the second round of workshops in June/July. There were no interviews in this round as I scheduled enough time within the workshop for the participants to talk about their respective data needs. I also modified the sequence of activities, as I alternated between general data literacy activities with databasic.io at one workshop and more context specific ones with the CWK dashboard at the next. I did this, as I explain in the next section, because some participants complained about working with data that they couldn't relate to in some of the databasic.io activities. Such modifications are not a deviation from the method and are in line with the design based research approach to learning (Sandoval & Bell, 2004) and the conversational approach of refining the workshops to the contextual needs of the participants. Participants came from the following organizations

1. **Historic Westside News** is an organization that produces the Historic Westside Newspaper is a monthly newsletter that focusses on giving voice to and connecting the different communities within the Historic Westside neighborhoods. Hailey was the only from this organization and was comfortable with manipulating data on the computer.
2. **The Conservancy at Historic Washington Park** works to conserve and promote the first park that was built for African Americans in South West Atlanta. Patricia

and Clara attended all of the workshops. Patricia seemed more comfortable interacting with data on the computer than Clara.

3. **Historic Westside Gardens** offers programs that promote home gardening with the goal of making residents of the Westside communities self-reliant and improve access to "healthy, fresh and affordable food". Casey was the only participant from this organization and was comfortable with manipulating data on the computer.
4. The **North West Ministerial Alliance** and the **North West Youth Power Inc** had one representative each at the workshops, David and Janet respectively, who often worked together to improve economic outcomes for youth in their neighborhoods. Both participants were not too comfortable with using data digitally.
5. **Raising Expectations** offers development programming that addresses the social, academic and mental needs of youth across several communities in Atlanta. Mary attended three workshops and Lillian dropped out after the first. Mary was comfortable manipulating data on the computer.
6. Roy was involved in a number of non-profit organizations but was there in his capacity as the chair of one of the **Neighborhood Planning Units** in Atlanta. He worked a lot with city officials and was comfortable manipulating data on the computer.
7. **WAWA (West Atlanta Watershed Alliance)** sent a second participant to this round of workshops. Kiara had an advanced degree (a PhD) and did quite a bit of work with data. She missed two of the workshops and agreed to host a makeup

workshop for herself and three other colleagues at WAWA. Of her two colleagues who attended the exploratory workshop, one was an intern who was studying at Georgia Tech and the other two were employees at the organization.

The table below gives an overview of the workshops I have just described. These workshops generated qualitative data in the form of written notes, pictures and audio recordings which I transcribed and analyzed using the Grounded Theory Method described previously.

Table 1: Overview of workshops

Workshop Type	**Exploratory Workshops**	**Round 1**	**Round 2**	**Make-up Workshop**
Date	March 2017	Feb/March 2018	June 2018	July 2018
Number of Workshops	2	4	4	1
Number of Participants in each workshop	5	7	9	4
Interviews	3	6	0	0

Activities	Session 1: Building data sculptures with Westside business data. Session 2: Introduction to visualization types and common pitfalls.	Session 1: Building data sculptures, Asking Good Questions with data Session 2: Sketch a Story with WordCounter and Try to ConvinceMe Session 3: Information hunt with CWK dashboard Session 4: Data 101: Data Visualization, Data Literacy and Storytelling	Session 1: Building data sculptures, Asking Good Questions with data Session 2: Information hunt with CWK dashboard Session 3: Sketch a Story with WordCounter and Remix a Visualization Session 4: Data 101: Data Visualization, Data Literacy and Storytelling	Session 1: Information hunt with CWK dashboard and Data 101: Data Visualization, Data Literacy and Storytelling

My analysis of these workshops resulted in narratives and themes about how community leaders understood data, the different ways they used data to tell stories, the kinds of data stories they valued, and the different barriers they faced when using data to tackle issues of community concern. These workshops also highlighted issues of data justice that needed to be considered when telling stories with data. I describe all of these themes and narratives in further detail in Chapters 5 & 6. I used the insights from these workshops to counter the injustices that were inherent in the dashboard and in my facilitation of the workshops. I did this by reimagining the Westside community's data infrastructure using the method of participatory infrastructuring as described next.

3.5 Step 3: Reimagining the Westside Community's Data Infrastructure

In reimagining the Westside community's data infrastructure, I adopted a participatory approach to infrastructuring, where the goal is to engage users in the design and

implementation of infrastructuring activities. Infrastructuring involves moving away from 'designing for use' to 'creating fertile ground to sustain a community of participants' (C. Le Dantec & DiSalvo, 2013). In describing participatory infrastructuring, Bødker et al. (2017) offer a vocabulary to use when addressing infrastructuring 'in the wild'. It emphasizes not only the traditional front-end process of PD that tend to be reported, like workshops, interviews and design activities, but also the back-end meetings, talks, chaos and technologies upon which infrastructuring builds. They differentiate between knotworks, which are more organic, fluid and temporary associations that evolve between participants who share an interest in common matters of concern. Networks tend to be built from the top-down are more stable and longer lasting. Such participatory infrastructuring is activated by creating new knotworks and expanding the networks, both horizontally and vertically. Here, vertical expansion involves including actors at different levels of political authority, while horizontal expansion involves a breadth of different actors.

The first step in the participatory infrastructuring of the Westside data infrastructure was to rethink the scenarios through which the community leaders would engage with the data infrastructure. Many community leaders expressed interest in creating and sharing data stories that highlighted their perspective, but not everyone was equally skilled at doing so. Additionally, these community leaders were engaged in a number of advocacy efforts and did not have the time needed to learn how to use new tools that would allow them to find data, analyze them and create compelling narratives. I describe additional constraints and design considerations in Chapter 7, but the need here was for a tool or a platform that offered enough scaffolding and documentation for novice users while also

allowing more advanced users the ability to create data stories of their choice. The tool needed to be collaborative so multiple users could combine their skills and work synchronously or asynchronously on a given data story. The suite of tools would need to be easy enough to install and manage for someone with limited computing skills, time and money. Access and security also needed to be considered, as only specific community leaders or their partners would be given access to create and modify the data.

Building an entirely new tool or platform from scratch was out of the question, as this would introduce some of the same issues around maintenance and sustainability that were concerns with the current dashboard. I looked for existing software tools and platforms that would support the identified data needs and concerns of the community leaders. I found that using a combination of Google's Data Studio, Google Drive and Google Groups met most of the identified needs. I describe the features of these platforms and the reasons I chose them in more detail in Chapter 7. In order to build the capacity to work with data, I complemented these platforms with the Communities Who Know (CWK) Data Ambassadors program.

3.5.1 The Communities Who Know Data Ambassadors Program

Before being dismantled in April of 2017, the WCA had been successful in serving the needs of the Westside communities thanks mainly to the engagement efforts of the Dean and her leadership team which included Sherri, Mackenzie and Katie. They had served as ambassadors for the WCA and developed a tremendous amount of good will and trust with all of the community leaders. Much of this good will and trust between the community and the university was lost when the WCA was dissolved as there was nobody to serve as a liaison between them (more about this in Chapters 4 & 5). The Dean

was the only familiar face that survived the WCA. While she continued her engagement and advocacy efforts through the new CWK brand and hired students to offer technical and programming support as needed, there weren't enough people to represent the CWK at all of the events and meetings happening in the Westside. This, as I describe in Chapter 5 & 6 contributed to the community leaders not knowing who at Georgia Tech they could contact to get errors in the data dashboard corrected.

The goal of the CWK ambassadors program was to identify individuals who had the time and interest to serve as data ambassadors for the CWK and others in the community. These ambassadors would work closely with CWK serve as liaisons between the university and the Westside for data related matters. The goal was to build enough expertise across the group of data ambassadors, so they could manage the entire data infrastructure themselves with minimal support from the university. To start off, I would work individually with these data ambassadors and support them as they acclimated themselves to the new data infrastructure I had put together. They would use the infrastructure to find data and create stories that were relevant to their concerns. My involvement and support would reduce over time, as the ambassadors would serve as points of contact for all data related questions to the rest of the Westside community.

I advertised the launch of this program by sending an email to a listserv that all organizations working in the Westside subscribed to. The launch took place at a local food bank and was attended by 14 community leaders, many of whom were also present at the previous data literacy workshops. The Dean did not have the funds to pay participants for their participation in the program, but we did offer them lunch in return for the time they spent with us that afternoon.

After the launch, I met individually with each of the participants who were interested in creating their own data stories. I took notes about the kinds of data they needed and the stories they were hoping to tell with them, details of which I present in Chapter 7. With this information in hand, I went about searching for relevant leads who could give me access to the required data. This process of contacting individuals and traversing the path to data was crucial as it revealed the underlying network through which data was made available. This process of working with the community leaders to learn about their needs, contacting relevant stakeholders for data, creating data stories with the bricolage of platforms I put together gave me insight into the human, material and organizational elements required to maintain and sustain a community data infrastructure. I used the Grounded Theory Method to categorize these insights into roles and activities for the identified infrastructural elements. I refined these categories using theories about the human infrastructure of information infrastructures (Bietz et al., 2010; Lee et al., 2006; Sambasivan & Smyth, 2010), which resulted in a set of guidelines that I used to describe how one can do justice to civic data when building data infrastructures for marginalized communities.

I draw attention to the different elements of the Westside's human data infrastructure in each of the next three chapters. These elements of the human infrastructure surfaced through the different stages of my interactions with the Westside community through the data dashboard. More specifically, these infrastructural elements came to light through the infrastructural breakdowns that I experienced while developing and engaging the community through the different elements of the dashboard. The story I tell in the next

few chapters is therefore a story of infrastructural breakdowns and their influence on the outcomes of my research.

In Chapter 4, I use the data dashboard as a window into the praxis, politics and infrastructural alignments involved in building the WCA Data Dashboard. This process revealed the *infrastructural bricolage* through which the Dean brought the different elements of the dashboard together. Chapter 5 describes the process I followed in organizing workshops for community leaders to further their data infrastructure literacy and perform an infrastructural inversion of the Westside's data infrastructure. The breakdowns I experienced in these workshops highlighted the barriers, challenges and opportunities of designing and using data dashboards for advocacy. I collated these findings into a set of workshop guidelines and local values of data equity that I outline in Chapter 5. These values build on the data justice scholarship I outlined previously as well as critical and feminist perspectives on data. These values guided the manner in which I countered the WCA dashboard and reimagined the Westside community's data infrastructure to be more collaborative, sustainable and scalable, which I describe in Chapter 6. I bring each of these individual contributions together in Chapter 7, which I then use to answer the overarching research question driving my dissertation.

CHAPTER 4. DASHBOARD AS A WINDOW INTO THE WESTSIDE DATA INFRASTRUCTURE

Data dashboards are data assemblages, which Kitchin, Lauriault, et al. (2015) define as "complex, politically-infused, socio- technical systems that, rather than reflecting cities, actively frame and produce them". The authors come to this conclusion by reflecting on the praxis and politics of building an urban data dashboard for the city of Dublin (Kitchin, Maalsen, et al., 2015). Building on this approach, I reflect on the praxis and politics of building the WCA data dashboard. In doing so, I focus on the term *build* that is in my research question, *What are the local values and infrastructural arrangements that are required to build, use and maintain equitable data infrastructures that enable communities to benefit from the publicizing of data through dashboards?*

The Dean envisioned the dashboard as a research project that would bring about equity in how data were produced and used in the Westside. This vision stemmed from her interactions with the community and her experience of using similar dashboards before coming to Georgia Tech, which she believed made data easily accessible. But this previous dashboard was used in a public university-advising context, which was completely different from the context in which the WCA dashboard was built. Even though the dashboard team claimed that the decisions to include specific modules were made based on input from the community leaders, many of the resulting features, as the rest of this chapter will show, were more a consequence of the politics and constraints with which the dashboard was being built. The dashboard infrastructure broke down in certain specific ways that impacted the manner in which its modules were built. I reflect

on these breakdowns, politics and constraints in the rest of this chapter and discuss how they shaped the dashboard and my research.

I first discuss how the dashboard came about and the process through which it was developed. I follow this with a detailed account of my involvement in building the Public Safety Module (PSM) of the dashboard. This lays the foundation for discussing the praxis politics of building the dashboard and the breakdowns that resulted from them. These breakdowns offer a window into the *infrastructural bricolage*, or the artful manner in which different elements of the dashboard infrastructure were integrated together. I describe this process by performing an infrastructural inversion of the dashboard and drawing attention to the specific ways in which it failed to meet the data equity goals it was founded on. I use these failures to argue for an alternate, justice-based approache through which such dashboards might be envisioned and built in the future. I conclude this chapter by outlining one such justice-based approach that came about through the infrastructural inversion of the data dashboard.

The benefit of doing such an infrastructural inversion of the dashboard is that is helps surface the people, processes, procedures, organizations, tools, facilities, and technologies which support its creation, maintenance and use. Knowledge of how these elements of the human infrastructure come together is the first step in answering my research question. In the process of unpacking the human infrastructural elements of the dashboard, I also draw attention to their inherent limitations, biases and challenge the assumption of objectivity that tends to be associated with them. More crucially, the infrastructural perspective helps bring attention to the always existing, relational and evolving nature of such data infrastructures and the critical manner in which each of its

taken for granted, forgotten and invisible socio-technical assemblages are connected (Bowker & Star, 1999).

4.1 Envisioning the WCA Data Dashboard

The Westside Communities Alliance Data Dashboard was a project borne out of individuals at the Ivan Alan College at Georgia Tech as well as the Westside communities. As such, I devote this section to profiling those individuals and communities that were involved in envisioning the dashboard. I choose to deanonymize the dashboard leadership team because their background and educational qualifications contribute to the narrative of how the dashboard came about and are needed to discuss the resulting power asymmetries.

I begin by sketching the professional profile of Dean Jacqueline Jones Royster, whose vision it was to build the Westside Communities Alliance Data Dashboard. Growing up in Greensboro, Georgia, Royster went to school at Spellman College in Atlanta and went on to do her M.A and D.A in English from the University of Michigan. Having spent several years Ohio State University as the Dean of the College of Humanities and then the Senior Vice Provost and Executive Dean of the Colleges of Arts and Sciences, Jacqueline Royster came to Georgia Tech in 2010 as the Dean of the Ivan Alan College of Liberal Arts. Dean Royster brought her interests in rhetorical studies, literacy studies, cultural studies, women's studies, and community engagement to initiate programs, awards and projects that strengthened the liberal arts program at Georgia Tech and its surrounding communities. Her knowledge of the issues that were prevalent in Westside Atlanta, which is just southwest of Georgia Tech, inspired projects that were related to preserving the history and building the assets of these neighborhoods.

Collectively referred to as the Historic Westside, these neighborhoods include English Avenue, Vine City, Bankhead, West Lake, Hunter Hills, Dixie Hills neighborhoods among others. Once a thriving middle-class community, desegregation, suburban flight, school closures and neglect from the city have resulted in these neighborhoods being riddled with issues that come with abandoned homes, joblessness, drugs and crime.

The sports arenas, Olympic parks and convention centers that were built in the 1990's in downtown Atlanta and promised to bring economic prosperity, increased investment, and vibrancy to the bordering neighborhoods of Vine City and English Avenue, failed to deliver. These massive infrastructure projects and their adjoining parking lots isolated the Westside neighborhoods from the Central Business District of Downtown Atlanta, which led to further devaluation of its homes and schools. Revitalization efforts have been planned over the years, and the recent opening of the Mercedes-Benz Stadium has brought a lot more attention and investment into these neighborhoods. Other major infrastructure projects like the Beltline Trail, Bellwood Quarry Park and the Rodney Cook Sr. Park, which are in different stages of implementation have also raised the profile of these neighborhoods

These developments have resulted in intense public debate and controversy as residents in the marginalized Westside communities have raised concerns of gentrification and development without displacement. Individual and private enterprises have tried to profit from the Westside's escalating real estate valuations, while non-profit organizations and neighboring universities are working with the city to rebuild blighted neighborhoods, make them safer, and formulate policies that will protect longtime residents from being forced out of their homes. Several individuals and organizations are

working with the city to foreground the voices and interests of the residents who have called Atlanta home for generations, which have not always been front and center.

As a neighbor and a key public institution, GT has also played its part in the revitalization of these neighborhoods. Residents though are unhappy and frustrated with some faculty and students who treat these neighborhoods as laboratories to carry out their academic experiments that last a semester or two, or until funding runs out. Once completed, residents rarely receive any document about the findings, or worse still, are given a report with details on how to 'fix' the neighborhood. This behavior is regrettable and condescending to residents who are well aware of the problems that plague their neighborhood and do not need outsiders to reiterate what they already know.

In order to help mediate and strengthen the relationship between the university and its neighboring residents, Dean Royster, with sponsorship support from the Ivan Alan College and the College of Architecture in 2011, setup a communications network called the Georgia Tech Westside Task Force. This was a network of faculty, staff, administrators, and students from across the Institute who were working with or interested in working with the Westside communities. The goal of this task force was to facilitate communication between its different entities and "to identify areas of mutual interest and commitment, areas where Georgia Tech can lend its intellectual resources and commitment to projects that support the health and vitality of our neighborhoods" (Royster et al., 2013). On April 23rd of 2011, this task force was renamed the Westside Communities Alliance (WCA). The WCA Executive Leadership Team included Dean Royster, GT Director of Community Relations Chris Burke, WCA Director Sheri Davis-Faulkner, and WCA Associate Director Mackenzie Madden. WCA functioned as an asset

for the community and was involved in the revitalization and development of Atlanta's Westside. This was modeled along the lines of the development approach taken by the Netter Center at the University of Pennsylvania, which uses strategies of academically based community service, university-assisted community schools and the anchor institution approach to engage students, researchers and the West Philadelphia communities in mutually beneficial partnerships (Harkavy, 2006). The WCA similarly established sustainable partnerships with residents and organizations by representing the community at public events, celebrating and advocating for its many tangible and intangible assets, and by organizing trainings and information sessions. They also operated a computer lab and a community newsletter that broadcasted upcoming events and resources that were relevant to the residents.

One consequence of all the interest in the Westside was that there was a lot of qualitative and quantitative data that was being generated about it. Students from neighboring universities and planning agencies were collecting data, proposing plans and generating reports about issues like economic development, subsidized housing, improved public transit, access to fresh food data among several others. As Dean Royster recalls,

> *What I found out that over the years, there has been lots of information generated, but most of it was not easily available to anyone. If you needed information, you had to go dig it up, or more than likely people were reinventing the same kinds of information over and over again, and it would end up on a shelf and not be useful for the next person.*

Dean Royster was frustrated that in spite of all this data being generated, it was inaccessible to those who needed it the most; residents and community leaders interested in using data to advocate for change in their neighborhoods. In her previous role at Ohio State University, Dean Royster had used dashboards to access data about students and their advising needs, as well as in her interaction with community organizations, non-profit and government agencies. Convinced that such a dashboard could be replicated in the Westside community context as well, Dean Royster proposed to create a data dashboard for community action that would allow community leaders, researchers, neighborhood associations and non-profits to have easy access to quantitative and qualitative datasets about the communities they lived and worked within. The decision to build the dashboard was one that was taken by a person in authority to realize her vision of what was the best course of action for the community and, as far as I could tell, not one that emerged from public consensus and deliberation, which is the ethos of Participatory Design. One must therefore question the need to build the dashboard in the first place, given that there was no evidence of the Westside community requesting such a dashboard or of a dashboard's success at achieving data equity for marginalized communities. Although valid and critical, it was too late for me to ask these questions as my involvement with the dashboard began in 2016, after the dashboard had already been launched. I consider these questions in Chapter 7, where I counter the WCA dashboard by reimagining what an equitable civic data infrastructure would look like. This chapter of my research is concerned with understanding the manner in which the WCA dashboard came together, which the WCA claim was a result of them responding to the needs expressed by the community, but in reality, was an attempt at implementing the

Dean's research objectives and a consequence of the sociotechnical constraints in which the dashboard was developed.

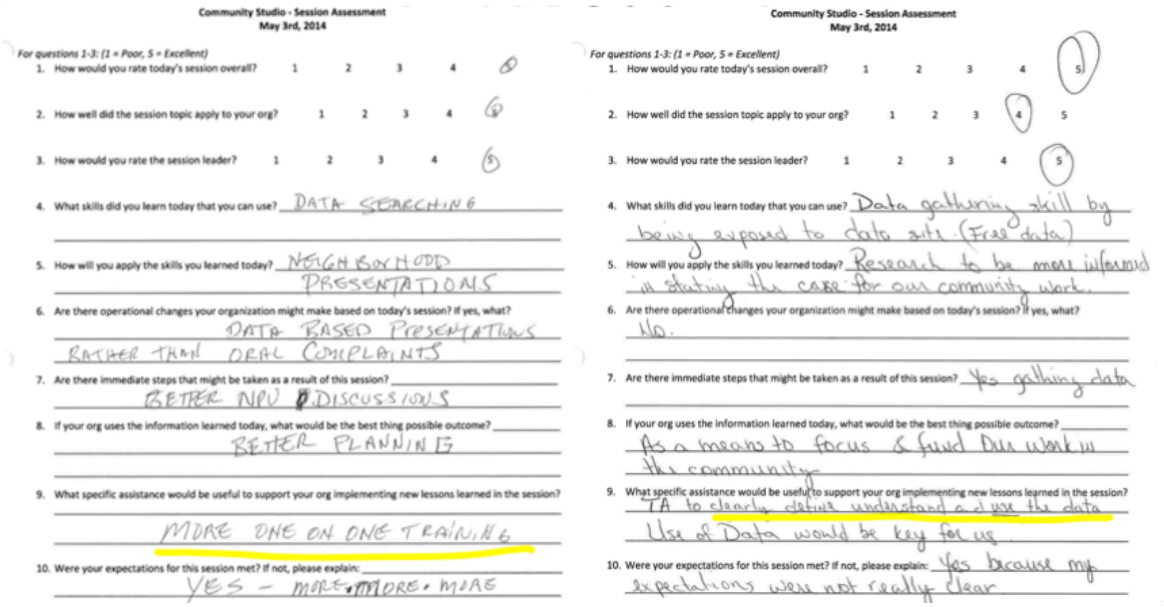
4.2 Building the WCA Data Dashboard

Katie O'Connell, a graduate of Georgia Tech's City and Regional Planning masters program was hired as the project manager for this envisioned dashboard. Having worked on Atlanta's Neighborhood Quality of Life & Health Dashboard (Botchwey et al., 2014) recently, Ms. O'Connell had just the skills and experience needed to realize Dean Royster's vision. Christened as the WCA Data Dashboard, this was positioned as the Dean's own research project and reflected her interests in historic preservation and community development and aimed to highlight the historical, geographic and economic significance of the Westside neighborhoods to Atlanta. It brought together previously inaccessible datasets from multiple city and state departments under a common platform that residents could use and analyze from the comfort of their homes.

The dashboard leadership team, which included Dean Royster, Ms. Davis-Faulkner, Ms. Madden and Ms. O'Connell, was keen on ensuring that the dashboard would meet the advocacy and reporting needs of the many organizations and community leaders that were serving the Westside neighborhoods. To realize this goal, they organized a Community Studio in the summer of 2014, which was a 12-week university-community course where academic experts led each weekly session in areas of housing, transit, planning/community development, communications, digital media, and more. The stated goal of the studio was to "engage multiple voices and perspectives in thinking about development and transit issues in support of sustainable development." [source: Ford

Foundation Report not available publicly]. Participants came from seventeen non-profit groups, which were compensated for the time their employees volunteered at the studio.

One of the twelve sessions focused on "Using Digital Data Tools", in which participants were introduced to websites where they could access census and health data from. The post-session questionnaire indicated that learning about "specific websites for helpful data" was beneficial. While some acknowledged the need for "more computer skills" and that the "use of data would be key for us", others requested "more one on one training" as well as "organizational development for the neighborhood" so they could better plan and communicate their goals. The session led to conversations about the usability of the existing tools and what they would like to see in a dashboard that would be tailor made for their needs.



Community Studio - Session Assessment
May 3rd, 2014

For questions 1-3: (1 = Poor, 5 = Excellent)

1. How would you rate today's session overall? 1 2 3 4 (5)
2. How well did the session topic apply to your org? 1 2 3 4 (5)
3. How would you rate the session leader? 1 2 3 4 (5)
4. What skills did you learn today that you can use? DATA SEARCHING
5. How will you apply the skills you learned today? NEIGHBORHOOD PRESENTATIONS
6. Are there operational changes your organization might make based on today's session? If yes, what? DATA BASED PRESENTATIONS RATHER THAN ORAL COMPLAINTS
7. Are there immediate steps that might be taken as a result of this session? BETTER INPUT DISCUSSIONS
8. If your org uses the information learned today, what would be the best thing possible outcome? BETTER PLANNING
9. What specific assistance would be useful to support your org implementing new lessons learned in the session? MORE ONE ON ONE TRAINING
10. Were your expectations for this session met? If not, please explain: YES - MORE MORE MORE

Community Studio - Session Assessment
May 3rd, 2014

For questions 1-3: (1 = Poor, 5 = Excellent)

1. How would you rate today's session overall? 1 2 3 4 (5)
2. How well did the session topic apply to your org? 1 2 (3) 4 5
3. How would you rate the session leader? 1 2 3 4 (5)
4. What skills did you learn today that you can use? Data gathering skill by being exposed to data sets (Free data)
5. How will you apply the skills you learned today? Research to be more informed in stating the case for our community work.
6. Are there operational changes your organization might make based on today's session? If yes, what? NO.
7. Are there immediate steps that might be taken as a result of this session? Yes gathering data
8. If your org uses the information learned today, what would be the best thing possible outcome? As a means to focus & fund our work in the community
9. What specific assistance would be useful to support your org implementing new lessons learned in the session? TA to clarify data, understand & use the data. Use of Data would be key for us.
10. Were your expectations for this session met? If not, please explain: Yes because my expectations were not really clear

Figure 14: Example post-session questionnaires filled out by one of the participants at the "Using Digital Data Tools" session.

Documentation regarding which of the dashboard modules were specifically requested for by the community and which were initiated by the dashboard leadership team does not exist. During one casual conversation, Ms. O'Connell mentioned that modules like Historic Data, Historic Timeline were added because that the leadership team saw a need for them and believed they would add value to the dashboard. Although not all of the modules were functional at launch in the spring of 2016, placeholders were included on the homepage (as shown in the figure below) to give visitors an idea of what they could expect from the dashboard in the future.

Figure 15: The first version of the WCA Data Dashboard.

Dean Royster recalls that when she initiated the dashboard project, not everyone within the university was convinced of the utility of a data dashboard for community engagement, which made university funding hard to come by. A grant from the Ford Foundation helped the dashboard project off the ground. Throughout the development process, the dashboard and its team survived on limited funds that the Dean could piece together from different sources, which is one reason why the WCA could not recruit an external organization to build out the dashboard. The WCA decided to approach faculty members and their students at GT to develop parts of the dashboard one module at a time.

What followed was a bricolage approach to building the dashboard, where experts were consulted, and different student teams were put together to focus on building out each of the modules. Ms. O'Connell, Ms. Davis-Faulkner and Ms. Madden were in charge of ensuring that these modules were built to a standard and incorporated a common look and feel, but many of the design decisions were also taken by the individual teams that developed the modules. I briefly describe the process and technologies that were used to build each of the modules next.

4.2.1 Community Profile

The community profile page contains information regarding the demographic makeup of the community, with data about racial breakup, housing, income, employment and health coming from the Census and the American Community Survey. Such data would be helpful for anyone who wants to know who lives in these neighborhoods and their socio-economic status. This page was first built using the D3 visualization library by a post-doc at the Center for Spatial Planning Analytics and Visualization center (formerly the Center for Geographic Information Systems) at GT. D3 is a powerful JavaScript visualization library with a steep learning curve, which the dashboard team found hard to maintain and update as they lacked the advanced technical skills required to do so. A GRA was therefore hired to migrate the page to Tableau, which is a drag and drop visualization platform and allows for quick creation and sharing of data analyses and visualizations. This move to Tableau resulted in a much more manageable dashboard that the team was able to update and maintain by themselves in spite of their limited programming expertise.

4.2.2 Education

Education is a huge concern among residents of the Westside community. The closing of public schools in the Westside neighborhoods, which came about after the desegregation and suburban flight is commonly known to be a catalyst in the downward spiral of the neighborhoods. The Washington Cluster Advocacy Group (WCAG)[15] is one of the many resident led groups that are fighting to keep their schools open (Bloom, 2016). The education page was meant to be a resource that would help groups like WCAG find the data they need to advocate for keeping their schools open. This page was also first built in D3, and then migrated to Tableau. The module has data and visuals that display dropout, enrollment, graduation rates and test score data either by Neighborhood Planning Unit (NPU) or Atlanta Public School (APS) district. Data for this module comes from the Georgia Governor's Office of Student Achievement.

4.2.3 Transportation

Freeways and roads that are traditionally meant to connect neighborhoods have instead resulted in economic and geographic isolation of the Westside neighborhoods from the rest of the city. The case of how Interstate 20 was laid down to racially segregate neighborhoods in the west side of Atlanta has been well documented (Kruse, 2019). Residents were demanding transit-oriented changes within their neighborhoods. They needed data about existing transit routes and how it correlated with different aspects of the census. A group of three students in a Data Visualization course developed the transportation page. The page, which was also built in D3 and allowed users to compare

[15] https://www.facebook.com/groups/wcagatl/

locations of MARTA (Metropolitan Atlanta Rapid Transit Authority) rails stations and bus stops with specific census data (population, age, race, education, household income, transportation to work, housing, and employment) on a map. Although rich in functionality, this module was slow, hard to update and not intuitive to use and was eventually dropped from the dashboard.

4.2.4 History

The Westside neighborhoods have a distinguished history that dates back to the Civil Rights Movement. With the recent demolition of two historic community churches, residents were concerned that their community's history was being erased. Residents have fond memories and stories to tell of growing up in a community that was once the home of Rev. Dr. Martin Luther King Jr. The neighborhoods have changed dramatically since then and it is this change that the dashboard team decided to capture and preserve through the historic data and historic timeline modules. The historic data page was first built in Google Charts, and was inspired by Hans Rosling's famous animated data stories. The page uses the same datasets from the community profile page but highlights the 'change over the years' narrative through animated visualizations. The visualizations on this page broke when Google made changes to their API, which meant that the GRA supporting the dashboard had to reprogram it using the Gapminder API developed by Rosling and his team at the Gapminder Foundation.

4.2.5 Historic Timeline

Starting right from the civil war, the Westside of Atlanta has been shaped by several key historical moments. The dashboard team wanted to have a page that showcased the

different civil and political events/policies that had a bearing on shaping the Westside of Atlanta. The historic timeline came out of this need to trace the current situation of the Westside neighborhoods to key policies, events and movements in history. Finding data for this timeline involved collating text and images from multiple websites and uploading them to Tiki-Toki, a service that allows one to create HTML based interactive timelines. The dashboard team hoped that such historic pages would be helpful to residents or researchers who are looking to learn more about the history events that shaped the Westside neighborhoods in Atlanta.

4.2.6 Resource Library

Dean Royster was frustrated that the same data and reports about neighborhood demographics, crime, education and other issues were being reproduced as there was no central location where they could be stored and accessed. The resource library was meant to be a repository for researchers, community leaders, city officials or other interested stakeholders to store and access data, resources and reports related to the Westside. Created using the open source web-publishing platform Omeka, the module archives historic as well as more recent reports, articles, development plans, maps and pictures related to the Westside neighborhoods.

4.2.7 Public Safety

The Public safety module was built by fellows in the Data Science for Social Good (DSSG) program to analyze the impact that specific public safety programs and features in the built environment were having on crime in the Westside neighborhoods. It includes modules that visualize crime and code violations data across the years in order to identify

patterns in the data. It also includes a map that helps spatially locate crimes, code violations and vacant houses with neighborhood assets like religious institution, schools, senior centers. Data for this module came from the Atlanta Police Department. The visualizations were built using Tableau. The map was built in CARTO and Leaflet (a JavaScript mapping library), since Tableau's mapping feature was very limited at the time. This module was developed further the following semester by teams of undergraduate students in the Computing4Good class. Realizing that fire was an additional public safety concern, one team worked on adding data from the Atlanta Fire Department to the map. They also built custom visualizations that allowed users to analyze how the number and types of fire incidents have changed over the years. I offer more details on the development of the Public Safety Module in the next section and use it as a window to reflect on the Westside community's data infrastructure.

4.3 Development of the Public Safety Module

Crime mapping has a rich tradition of scholarship (Eck et al., 2005; Lodha & Verma, 2000; Santos, 2016; Tabangin et al., 2008) and most major cities have crime maps that depict the number and pattern of crimes in its different neighborhoods. Some have gone beyond this traditional mapping of crimes to map the origins of those incarcerated to reveal patterns that link poverty, racial segregation, and incarceration with spatial features of cities (Kurgan, 2013).

The idea for the Public Safety module was a result of the blight that the WCA witnessed firsthand through its many engagements in the Westside. Many homes in the Westside neighborhoods remained abandoned and had become centers for crime and illicit activities. Crimes, their relationship to features in the built environment and public

safety were top concerns for the Westside residents. In addition to their own personal safety, residents were also concerned that being perceived as an unsafe neighborhood also prevented visitors and businesses from coming to them. Newspaper articles and popular media tended to highlight the negative stories of crime, addiction and joblessness, which further contributed to the public's fear of these neighborhoods and ended up alienating them from the rest of Atlanta (Whitney, 2011). Residents and the WCA were trying to change the narrative that gets told about their community, to one that highlights its assets and stories of change rather than just the negative aspects.

This was the brief our DSSG team got from the WCA when we started to work on the Public Safety module (PSM) of the dashboard in the summer of 2016. Our team included an undergraduate student in Industrial Systems Engineering, and a graduate student, who was close to graduating with a PhD in Physics. Both of these students had strong skills in math and data analysis, but not as much with web programming in HTML/CSS and JavaScript. All DSSG fellow went through a one-day boot camp in which we learnt the basics of web programming. The graduate student who organized the boot camp was also available to consult with the teams throughout the duration of the fellowship. Each team also had a faculty mentor who met with them weekly and guided them through the project. Teams also published a weekly blog in which we documented our thinking, findings and progress so we could reflect on it later.

In addition to the faculty member, our team also met weekly with the dashboard team (Mackenzie, Katie and sometimes Sherri) in which we discussed ideas and clarified questions we had about the purpose and scope of the Public Safety module. The WCA was the client in this relationship, who was working with us to design the module for its

users in the Westside. We got most of our requirements for the module from the WCA because they claimed they had already engaged with the Westside and were aware of their needs. We were encouraged to meet with the community at public events to get a firsthand experience of the issues they were facing. We did this by attending Neighborhood Planning Unit (NPU) meetings, public safety committee meetings and meeting with public safety officers in the community.

Our participation in these events and discussions with the WCA led us to shortlist three use cases for the public safety module to support. One that was requested by the leadership team, which they insisted came out of their previous interactions with the community leaders, was the ability to locate crimes, code violations and vacant houses on a map of the Westside neighborhoods. This would allow individuals to visually determine if the number of code violations and crimes were in any way related to the number of vacant houses in specific geographic areas. Additionally, the map would include the ability to correlate crime and code violations with specific aspects of the built environment like vacant homes, parks, schools and churches, or even with public safety programs in these neighborhoods.

A second feature requested by both the leadership team and the community stakeholders we spoke with was the ability to report incidents, so their investigation could be tracked via the dashboard's interface. Such a feature would require an interface where users could report a public safety concern or incident, which the dashboard system should then be able to verify and track as it was investigated by the relevant authorities. Many in the community believed that data from official sources did not reveal the entire picture of how many crimes took place in their neighborhoods and what was being done to

investigate them. We debated this feature extensively and decided not to pursue it. This was primarily because such detailed data about incident investigation was not available from the APD. Additionally, our team did not have the programming expertise needed to build out such an authentication system in the ten-week time frame of the fellowship.

A third feature, not specifically requested, but one that our team ended up building was the ability to correlate crimes with other census parameters like education, race, age, transportation, income, housing etc. We built this feature as an experiment, to know how the census data was correlated with crime and the built environment in the Westside. In hindsight, this was problematic because the module was driven not by the needs of the community, but by our desire to experiment with the data that was available. The correlations module ended up being the most controversial and did not make it to the final version of the PSM. I discuss the reasons for this resistance, along with other ways in which the PSM was shaped by the community in the next section.

We used Tableau to build the visualizations in the PSM as this was used in other dashboard modules as well. We learnt how to use Tableau as none of us had experience with it. The mapping features in Tableau were limited at the time, which is why we used web-based mapping libraries like Leaflet and CARTO to build the map. We have described the software architecture of this module in further detail in O'Connell et al. (2016).

Part of our data to build our system came through the Freedom of Information Act we filed with the Atlanta Police Department (APD). We also met with one of the officers at the APD's Office of Code Enforcement, who shared their data with us, and answered questions we had about data provenance, cleaning and use. We downloaded census data

from Neighborhood Nexus' website because it was preformatted to Atlanta's Neighborhood Planning Units (NPU's), which was useful for our analysis. And lastly, data about the different assets within the neighborhoods, like schools, parks and religious organizations was downloaded from the Atlanta Regional Commission's website.

4.4 The Praxis and Politics in Building the Public Safety Module of the WCA Data Dashboard

Kitchin et al. (2015) use the terms 'praxis and politics' to refer to the everyday work that goes into building urban dashboards and how this work influenced by the politics the environment within which it is embedded. They claim that data dashboards are sociotechnical assemblages that evolve relationally and contingently based on the capacities of its actors and actants. This evolution was visible in the WCA dashboard as well. My participation as an ethnographer and a developer of the public safety module revealed how it shaped, and was in turn shaped by the conventions, beliefs, knowledge, expertise, praxis and politics, both within our team and the stakeholders our team was working with. We witnessed this shaping at three levels; the *interface* level, the *community* level and the *infrastructural* level.

Consider the name 'Public Safety', which was chosen to steer clear of the negative connotations associated with names such as Crime Map or Crime Statistics. The goal of this *interface* was to not only highlight the areas where crimes and code violations were a concern, but also draw attention to public safety programs and assets these communities were using to combat crime in their neighborhoods. This was the reason our map included additional layers for religious institutions, senior living centers, parks and law enforcement centers that highlighted the assets of the community. Specific layers could

be enabled to show the number of crimes and code violations that occurred in each year or month around these community assets. Such a map would be helpful for churches and organizations that are organizing public safety programs, so they can visually determine if their efforts are resulting in a reduction of crime around their locations. The map could also be used as evidence to show external funders and media organizations the difference that these programs were making in the community.

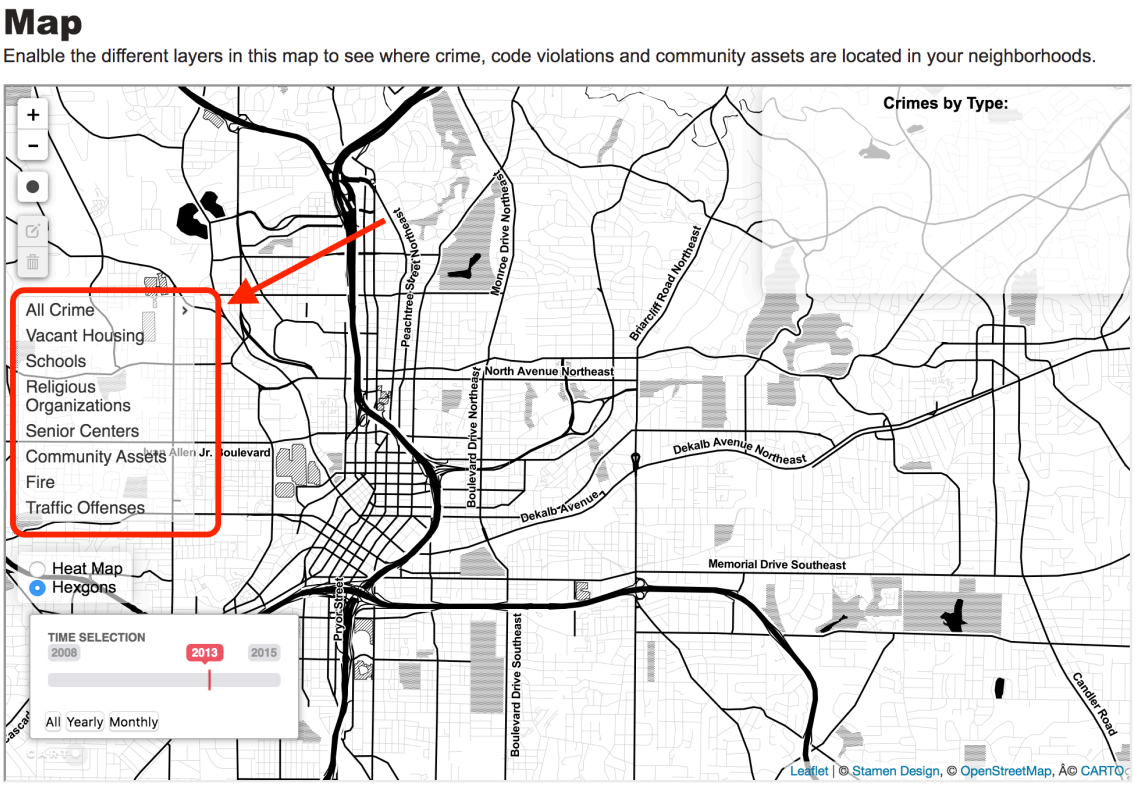


Figure 16: Map of the public safety module showing different layers for crime and community assets.

We experienced resistance from the *community* when presenting the experimental correlations module. This module, as shown in the figure below, allowed one to correlate crimes with other census parameters like education, race, age, transportation, income, housing etc. This was problematic, because the visuals revealed relationships between

factors that had no evidence of being related to each other. Violent crimes, for instance, were positively correlated with the number of senior citizens in a neighborhood, but this said nothing about why or how they are related. At the mid-semester presentations that we made to the campus and Westside community, we found ourselves constantly having to clarify that correlation did not imply causation. The WCA insisted that they did not want these correlations to be made public on the dashboard. Given the possible controversies and negative connotations such a misinterpretation could have for the community, especially one that is consistently portrayed as unsafe and dangerous, the dashboard we decided to shelve the correlations part of the PSM.

Figure 17: Comparison of different correlations in the Public Safety Module

Another point of contention was the overall branding and the presence of the original WCA dashboard. The fact that it was (at the time) hosted on a GT domain was controversial, as it raised questions about who owned its data. If the data did indeed belong to the community, why did it need to be hosted on a website under the GT domain? Why could it not be under an organizationally neutral domain that suggested the

data was owned by the community? This *infrastructural* issue was in some way resolved when the WCA ceased to exist.

I first leaned about the WCA being restructured from Katie in April of 2017. The dashboard was taken offline without any warning and the WCA staff were asked to cease all activities as the Dean and GT figured out the future of the WCA. The fact that Dean Royster had failed to secure substantial funding to support the WCA and was relying on piecemeal funding she put together from multiple sources might have been one nail in WCA's coffin. Some faculty and staff I spoke with informally suggested that the university leadership still did not see value in the dashboard, which is why they chose not to fund or adopt it. When I asked the president at a breakfast gathering why the university was not as supportive to building relationships with the Westside through projects like the dashboard, I was reminded of how complicated the university-community relationship was and that the university was investing a lot more money into those neighborhoods through other programs. When I emailed Dean Royster on May 1st expressing my interest to continue supporting the dashboard, I was told that the WCA was "in some sort of transition" and that the Dean was unsure of what the future would hold for the WCA.

Eventually it was decided that WCA would cease to exist and its staff would be absorbed by other academic units on campus. Although GT did not value the dashboard, they did value the connections and good will that the WCA had built with the Westside communities over the years. Mackenzie was absorbed into the Office of Community Relations so she could continue the work she was doing under the WCA, but now in an office that reported directly to the President. The hope was that this would bring more

visibility to the efforts GT was making in building community relationships. This was an uncomfortable fit, I was told, because the Office of Community Relations prioritized building relationships with state and federal agencies rather than its neighboring communities. Expectedly, Mackenzie's new work arrangement did not last long and she transitioned out of GT by the end of that academic year.

In spite of the disbanding of the WCA and the reshuffling that took place, Dean Royster retained control of the dashboard. She took it offline for a few months till she could figure out the organizational and monetary structure to support it. The restructured dashboard was launched at a community event under the Communities Who Know Inc. banner, which also included other archival and youth projects that the Dean was leading. Dean Royster decided to place the dashboard under a non-profit as it opened up other sources of funding that were not available when associated with the GT brand. The dashboard went through some substantial polishing over the fall of 2017, as it was decided to retain only the modules which were completely functional.

This restructuring of the dashboard also had a significant impact on its usage as reflected in the web analytics. Before it was taken offline in May of 2017, the WCA dashboard had an average of 243 users per month. This number was averaging at 28 users per month for the CWK dashboard, which went online in October of 2017. There might be several explanations for this dramatic drop in the number of users. For one, the leadership team was actively promoting the dashboard at community events and neighboring schools and universities, which did not resume when the dashboard was relaunched. The grand vision of being a one-stop shop for community related data, which is evidenced by the number of modules that were initially planned, had to contend with

administrative, economic, social, political and cultural contingencies that ended up changing what the dashboard could offer. To put it differently, the breakdown of the dashboard's human infrastructure led to the dashboard being taken offline and unable to consistently meet the needs of the community.

4.5 The Dashboard as a Window into the Westside Data Infrastructure

The dashboard has changed considerably since it was first launched in the spring of 2016. What started off as a typical computing for good project, where students built dashboard modules for the community, quickly turned into an act of bricolage where multiple software platforms were seamed together in order to achieve the task at hand. The dashboard did not come about according to a set plan but had to grow in response to the environment in which it was being built. This idea of growing like a plant in response to its surroundings is what some authors have argued is an essential characteristic of information infrastructures (P. N. Edwards et al., 2007; Jackson et al., 2007).

The dashboard grew from a stable base of datasets released by federal and state agencies like the Census Bureau, Georgia Department of Education, Atlanta Police and Fire Departments. Other cities, counties and non-profits also use datasets from these sources to measure their progress towards specific goals. Specific technologies used to build these indicators also vary, but are held together by legal, technological and institutional standards and conventions. The .csv (comma separated value) file format, for instance, is now almost universal. Most agencies that distribute data tend to have a .csv file available for their users to download and use within their development platforms. When CARTO has to display co-ordinates on Google Maps, data from these .csv files are translated into a different JSON (JavaScript Object Notation) standard before they can be

exchanged between web apps. These standards are examples of *gateways* that allow heterogenous systems within an infrastructure to seamlessly connect to each other (P. N. Edwards et al., 2007). Such gateways are not imposed in a top down fashion but develop into a standard as they are adopted by a wider range of systems and developers. These datasets, institutions, standards, and conventions form the *installed base* on which data infrastructures are constructed and contribute to the transparent fashion in which they operate.

Further commenting on the idea of an installed base, Hanseth & Monteiro (1998) consider infrastructures as 'always already existing', in that they are never built from scratch and are always embedded in or replace parts of existing infrastructures. The dashboard infrastructure, like the Internet, World Wide Web infrastructures on which it is based, did not form overnight, but came about through a gradual accretion of technological, social, legal and economic enablers that made such infrastructures feasible. The different modules started off as prototypes within class projects and fellowships and made use of existing components to demonstrate a proof of concept. These individual modules were then consolidated through the convergence of multiple actors and networks across multiple sites and over time.

Infrastructures also inherit the limitations and capabilities of its installed base (Marttila & Botero, 2017). The platforms that the WCA chose to build out its dashboard also had implications for the kinds of GRA's it hired and the work they had to do in developing and maintaining the different modules. Although the WCA was successful in utilizing the scores of bright engineering talent available at Georgia Tech, the process wasn't always easy. Not all students had equal programming capabilities or adhered to strict coding and

documentation standards. As newer GRA's tried to pick up from where the previous ones left off, there is often some knowledge that does not get transferred as well as it should. Students' schedules change every semester and as new ones join, sensitizing them to the predicaments of the community and knowledge of how to navigate the idiosyncrasies of the dashboard and university-community relationship is always a challenge, and something the WCA always played catch up with. This means that the dashboard is constantly trying to keep up with the sociotechnical changes in its environment.

Conflict and negotiation featured constantly during the infrastructuring process, as there was rarely ever one correct solution to problems faced by the communities (P. Edwards et al., 2009). The university not willing to support the development of the dashboard, skepticism on behalf of the community and the various socio-technical hurdles that the dashboard team had to endure are all examples of *reverse salients;* the legal, political, cultural, social or technical issues that need to be solved in order for the infrastructure to develop and grow (P. N. Edwards et al., 2007). Part of infrastructuring the dashboard asked that the leadership team present and promote the dashboard at local events and community meetings. As a communications network, the WCA built trust through its operations by always highlighting and celebrating the voice of the community. In addition to hosting sessions like the Community Studio which I described earlier, the WCA also organized events like *WCA Salutes Westside* and *WCA Community Development Symposium & Awards Luncheon* that recognized individuals and programs who had made a difference to the Westside. They also organized workshops at Clark Atlanta University, Spellman College and Georgia State University that were meant to showcase the capabilities of the dashboard.

The WCA setup a computer lab where software trainings and information screenings around housing, food and health benefits were periodically organized. The lab was an important part of the community's information infrastructure as it allowed residents free access to computers, WIFI and benefits screenings. Its name changed from the 'Westside Computer Lab' to the 'Westside Resource Center' to reflect the center's role as more than just a technology hub but a community resource center in general. The center served a space for those who did not have easy access to technology to learn about and apply for the different benefit programs related to housing, health, job training etc. that the city and other organizations had put in place. It also served as a space where dashboard demos and trainings were held.

This form of convergence of social and technical systems also necessitates interfaces with human behaviors, standards and practices. As Star et al. (2003) claim, convergence is a concept which expresses "the double process by which information artifacts and social worlds are fitted into each other and come together". This convergence of different systems implies that infrastructures simultaneously operate at multiple levels (Larkin, 2013). In addition to being a tool that makes data available to a community, the different modules of the dashboard hope to recognize the social, cultural and historic significance of the Westside communities and its residents. The Dean is continuing to hire students to add data and build out additional modules in the dashboard. By highlighting the community's assets rather than its shortcomings, the dashboard is trying to change the story of the Westside to that of progress and development rather than dilapidation, neglect and isolation. Funding for all of these efforts continue to come in piecemeal fashion.

My ethnography of the dashboard served as a window into the praxis and politics of infrastructuring a dashboard for community engagement and revealed how the underlying infrastructural elements came together. The process, as I will argue below, was a form of *infrastructural bricolage*, in which *Community Champions* like Dean Royster relied on *Transient Workers* like students, faculty and staff at GT to build out the dashboard. The dashboard was also influenced by the *Institutional* policies and *Organizational* structures within which it was built. These represent elements of the human infrastructure that were most relevant to this phase of infrastructuring the dashboard. I summarize these actors in the table below and discuss them further in the following section. I will add to this table over the next three chapters as I encounter more of these infrastructural elements.

Table 2: Elements of the Human Infrastructure

Infrastructural Elements	**Humans that Put the Infrastructural Elements Together**	**Type of Human Infrastructure**
Dashboard, funding, organizational structure, management	Dean	Community Champions
Individual dashboard modules, interactions with the community, software platforms	Graduate Research Assistants, Other Researchers, Faculty and Staff	Transient Workers
Communications network (newsletter, website, dashboard), programs, workshops, data/dashboard/infrastructure maintenance, Westside Resource Center	Westside Communities Alliance/ Communities Who Know Inc., Other Non-profit Organizations	Organizational
Data, policies, protocol, policies, funding, philanthropy, tax regimes, laws, licenses, jobs	Political and Educational Institutions. Schools, Colleges, City, County, State and Federal Agencies	Institutional

4.5.1 Infrastructuring as Bricolage and Artful Integration

Levi-Strauss (1966) describes bricolage as the practice of making with what is materially or culturally available rather than building systems from scratch. It embodies a way of thinking that is fundamentally different from engineering. Where the engineer starts off by procuring the raw materials required to get the job done, the bricoleur works with a fixed set of tools and materials that are available at hand.

Ciborra (2007) traces bricolage to its French origins, in which it implies a way of tinkering. While the English meaning can often be negatively associated with being unprofessional, trying to stay afloat or doing too little too late, Ciborra reminds us of the several virtues of operating through bricolage. First, they are highly situated, relational and are deeply rooted in the personal and collective skill and experience of the individual or their community of practice. It takes a certain amount of skill and resourcefulness to be able to intervene in unknown situations where several sociotechnical forces are at play.

This concept of bricolage has been used to analyze how children learn computer programming by associating ideas with one another rather than following a top down process (Turkle & Papert, n.d.), how interaction designers can practice bricolage to take the field forward (Vallgårda & Fernaeus, 2015), how software can be developed for small enterprises (Büscher et al., 2001). In the context of information infrastructure projects in developing countries, Ali & Bailur (2007) have described how normative approaches to information sharing need to be tailored to the specific needs of their communities in order to be relevant and sustainable. In each of these instances, the authors describe projects where the students/designers/architects/workers drifted away

from what is considered the traditional practice in their field to one that is designed to meet their immediate needs and is contextually situated.

Erickson & Sawyer (2019) use the term *infrastructural bricoleur* when describing how mobile and independent knowledge workers, like real estate agents, retrofit their work environments in specific ways to get their work done. They refer to the human and non-human elements of the infrastructure that need to be brought into alignment in order to accomplish their task and overcome moments of infrastructural breakdown or impediment that show up repeatedly. These knowledge workers exhibit an *infrastructural competence*, which Sawyer et al. (2019) define as

> *an individual's use-oriented relationship with infrastructures that combines social abilities, goal-orientedness, and leveraging of digital and material resources in a way that enables one to generate a functional, operable, and patterned or routinized (while also being personalized), set of sociotechnical practices to accomplish a necessary task or set of tasks.*

Vertesi (2014, p. 278) similarly describes how scientists collaborating from geographically distant locations work creatively with and across infrastructural seams by *artfully integrating* (L. Suchman, 2002) various sociotechnical systems like power outlets, teleconference software and communication services into a "novel patchwork from which new arrangements, alignments, and actions can fruitfully emerge". Such *artful infrastructuring* of information systems has also been described by Karasti & Syrjänen (2004) in the context of scientific research networks as well as dog breeding enthusiasts, both of which deploy decentralized, bottom up processes that bridge the gap between technology use and design in the development process.

My intention in reviewing terms such as infrastructural bricolage, infrastructural competence, and artful infrastructuring is to draw attention to the manner in which Dean Royster and others at the WCA operated in bringing the dashboard together. The term bricolage can also refer to DIY or 'do-it-yourself' mode of operation where individuals rely on their specific communities of practice when getting work like house maintenance, gardening, or repair work done. But the concept of bricolage and infrastructural competence I have descried above is distinct from *DIY Infrastructures*, which is a term Lukens (2013) uses to describe how individuals or groups bring about systemic change by designing to counter the power of existing sociotechnical infrastructures. The work I am describing was done not to counter the monopoly of existing data infrastructures, but to build on their installed bases and gateways (P. N. Edwards et al., 2007) in order to produce something novel.

Dean Royster operated as a bricoleur and exhibited infrastructural competence in the process of bringing the data dashboard to fruition. When she envisioned a data dashboard for community engagement in 2011, she was responding to a need that she thought was relevant at the time, around the lack of access to usable data for the Westside community residents. In addition to having the foresight to build such a tool, the Dean had the power to artfully integrate multiple resources to make it happen. Establishing the WCA so she can hire personnel to serve as community liaisons and do the technical work of building the dashboard was her way of infrastructuring the Westside data infrastructure.

Once the funding sources dried up, Dean Royster had no choice but to dismantle the WCA and figure out alternate ways she could keep the dashboard afloat. Her response was to form a non-profit organization called Communities Who Know Inc. that would

serve as an umbrella organization for the dashboard and other community projects she was leading. Development on the dashboard is ongoing, as Dean Royster continues to apply for funding so she can support GRA's who can update the data and the layout of the different dashboard modules. Dean Royster operated as a bricoleur who used the resources and tools at her disposal to respond to the changing social, material, economic and technological issues that surfaced during the development of the dashboard. The Dean represents a category of human infrastructure called *Community Champions*, who are actors that can work across spheres of influence and have the power and resources to influence change. Other actors within such a category could include mayors, presidents, faculty, elected officials or other individuals in positions of power and influence to bring about sociotechnical change.

The role of the *keystone species* in data ecosystems are similar to the role that *Community Champions* play in setting up and enabling data infrastructures. *Keystone species* (Nardi & O'Day, 1999) are those who create value by bridging institutional boundaries and translating data for specific uses. These "keystone species are enablers, not necessarily drivers in the ecosystem; they can be useful but they are not essential to the sustained functioning of an ecosystem" (van Schalkwyk et al., 2016, p. 77). But while the metaphor of an ecosystem helps draw attention to the diversity, complexity and evolution of the people, practices, technologies, and values in a local data setting, they do not necessarily emphasize the power differentials and injustices that commonly accompany open government data initiatives (J. Johnson, 2014). Additionally, it also does not attribute equal agency to the human and non-human actors in the network,

which is something that the Actor Network and Information Infrastructure theories offer more guidance on.

Several *Institutional* and *Organizational* actors also influenced the trajectory of the dashboard and its infrastructure. As the owner and initiator of the dashboard, Georgia Tech and its several departments, faculty, staff and students influenced the manner in which the infrastructural elements of the dashboard came together. They were in turn guided by policies, issues, funding decisions taken by city, state, federal and private agencies which also had a stake in the Westside. State and Federal agencies are typically in charge of determining policies and setting the standards for what gets done at the city and county level. Such organizations and institutional policies have a varying amount of influence depending on the issue at hand and other socio-political considerations.

Transient Workers like GT students and the WCA staff also played a crucial role in infrastructuring the dashboard. Students were responsible for building out individual modules of the dashboard under the guidance of the WCA staff. The WCA in turn made sure that they were abreast of the ongoing issues on the Westside so they could serve them appropriately, either through the dashboard or any of its other services.

In spite of these efforts, the WCA was unable to foresee and keep up with the changing sociotechnical context within which the dashboard was deployed. The WCA's socially good intention of making data accessible on a public platform proved to be inadequate as it ignored the contingencies within which dashboard infrastructures are developed and used, which are further complicated by resource constraints that are typical within universities and minoritized communities.

The dashboard team overcame its resource constraints by tapping into its most abundant resource, its university students. Many universities offer programs like Computing4Good, AI for Good and Data Science for Good, which operate by partnering university researchers and students with organizations that are working towards improving the social wellbeing of their communities. Community organizations and their employees serve as domain experts, while university students and researchers bring the technical expertise required to build the necessary computational systems. Such partnerships can be a form of techno-utopianism that privilege values like efficiency, scale and speed of the technical artifacts, which are misaligned with the values and goals of the community. These partnerships are complicated by the fact that they have to contend with a sociotechnical system in which needs, priorities, and resources are constantly changing.

One consequence of this constantly changing sociotechnical system was that the WCA struggled with the maintenance and upkeep of the data dashboard. All of the modules that were built in D3 had to be redesigned in Tableau because the dashboard leadership team found them hard to maintain. Even the visualizations and maps our team built into the Public Safety module had customized data analysis code that was not easily decipherable to the students from the Computing4Good class who continued to work on it the following semester. This is one of the reasons that data in this module has not been updated to include more recent datasets from the police and fire departments. This issue of maintenance and sustainability for technologies deployed within marginalized communities has been a topic of concern within the HCI community (Gonzales, 2017; Huh et al., 2010; Jackson et al., 2012; Rosner et al., 2013; N. Taylor et al., 2013).

Scholars are unanimous in calling for more attention to be paid to how marginalized communities can continue to have access to technologies that have been developed and deployed for them. This includes considering the financial, technical, material, relational resources and alternative design methods that are needed to design, use, maintain and sustain these technologies. This points to the need for an alternate approach to infrastructuring dashboards, one that is different from designing for social good.

4.5.2 An Alternate Approach to Infrastructuring Data Dashboards

If the dashboard is to serve as a resource for data advocacy, equity or justice, it's availability cannot be subject to institutional power struggles and academic funding cycles. If data is a right, its unavailability through the dashboard is a form of social injustice that needs to be overcome. Structuring the dashboard assemblage so it relies on more of the community's distributed data infrastructure would make it less reliant on the vagaries of a single institution. Such an infrastructure would include the humans, materials, technologies and organizations that can support the collection, curation, cleaning, control, collaboration and creation of data. A social justice approach to designing such a dashboard infrastructure would require that we reflect on the context, its politics, ethics and injustices, look for what is already working at the community level before designing new solutions, and work towards sustainable, community-led, and controlled outcomes that prioritize design's impact on the community over the intentions of the designer (Costanza-Chock, 2018; D'Ignazio & Klein, 2020; Dombrowski et al., 2016).

Balestrini et al. (2017) recommend starting with matters of concern (Latour, 2004), building technical capacity and data literacy within the communities, negotiating data

ownership and governance strategies and finally creating growth opportunities for local entrepreneurs. The role they suggest for the designer/researcher in this framework, is that of a helper, someone who can explain and help put out fires rather than control or manage. This implies that their role in the infrastructuring process is relegated to that of a facilitator who is committed to surfacing community conflict, while being reflexive of their own ethics and politics of the situation they are designing for (Costanza-Chock, 2018; Dencik et al., 2019; Walker, 2018). Such a role relegation, in addition to building the community's data literacy through a collectively owned and controlled data infrastructure might lead to more socially just outcomes for the communities that the dashboard is meant to serve.

I explore what such a social justice orientation to infrastructuring the data dashboard would look like in the rest of this dissertation. This chapter has laid the foundation for such an exploration. I started off by aligning myself with Kitchin, Lauriault, et al's. (2015) interpretation of data dashboards as sociotechnical assemblages. I conducted an ethnography of the WCA data dashboard that was inspired by a similar ethnography of the Dublin Dashboard (Kitchin, Maalsen, et al., 2015). In addition to revealing the praxis and politics of infrastructuring the WCA data dashboard, this ethnography also revealed the artful ways in which the different elements of the human infrastructure were brought together in order to function as a whole. In trying to realize her vision of an equitable civic data dashboard, Dean Royster worked as an infrastructural bricoleur who used her power and influence to bring other individuals, institutions, and organizations to align with her cause. I classified these elements of the human infrastructure as *Community Champions, Transient Workers, Institutions and Organizations* (Table 2). The

breakdowns that I experienced point to a justice-based approaches that one could explore when infrastructuring such dashboards for marginalized communities. Such a justice-based approach, which I discuss in Chapters 5 & 6 would not be based on universal values of justice, but on local values of justice that are grounded in the experiences of the community one is designing for. The next chapter describes my approach to determining these local values of justice and reimagining the Westside community's civic data infrastructure along those values.

This chapter described the process I followed to perform an infrastructural inversion of the WCA dashboard that hoped to meet the data equity needs of the Westside communities in Atlanta. Following in the footsteps of Kitchin, Maalsen, et al. (2015) I discussed the praxis and politics involved in conceptualizing and building data assemblages like indicator data dashboards for marginalized communities. The dashboard further served as a window into the infrastructural breakdowns and evolution that took place within the Westside's data infrastructure. In using a data assemblage like the WCA dashboard as a window into the infrastructural inversion of the Westside community's data infrastructure, I situate the study of data dashboards within the wider field of Information Infrastructures. The sociotechnical unpacking of the dashboard offers evidence of how the concept of *infrastructural bricolage*, in which the Dean worked as an *infrastructural bricoleur* and attempted to intervene in the Westside through the artful integration of data through a variety of sociotechnical platforms. Thus, the concept of bricolage, which has previously been discussed in the domain of information systems (C. Ciborra, 2007), computer programming (Turkle & Papert, n.d.) and workplace practices (Erickson & Sawyer, 2019) can also be applied to the design of data dashboards and their

underlying infrastructures. This would be useful for designers, researchers and practitioners in HCI and CSCW who are looking to build similar data dashboards and infrastructures for their communities. Finally, this first phase of my research points to the inadequacy of data equity initiatives that are motivated by the idea of social good and not grounded in the community's local values of justice. Based on these results, I align myself with other scholars of data and design justice (Bardzell, 2010b; Costanza-Chock, 2020; D'Ignazio & Klein, 2020; Dombrowski et al., 2016; L. Taylor, 2017) to argue for an approach in which data equity initiatives are driven by local conceptualization of justice rather than a universal one.

CHAPTER 5. DATA LITERACY WORKSHOPS FOR DESIGN AND INFRASTRUCTURAL INVERSION

The WCA had organized focus groups and demos of the dashboard at several gatherings around the city and were keen on organizing data literacy workshops as they considered it a part of infrastructuring the dashboard within the Westside. I approached them in the Spring of 2017 and proposed organizing these workshops for them. We hoped that these workshops would help ensure that the dashboard and its data could be used to their potential. This chapter describes my approach to organizing the data literacy workshops, the activities that ensued, the observations I made as well as my contributions.

These workshops have resulted in two scholarly contributions. First, I offer a set of guidelines that researchers and practitioners can use when trying to build data infrastructure literacy with their communities. Data infrastructure literacy, as proposed by Gray et al. (2018), is distinct from data literacy, in that it focuses on surfacing and as a result intervening in the broader sociotechnical infrastructures within which data are created and used. This notion of data infrastructure literacy aligns with my own goal of using the WCA dashboard as a window into rethinking the Westside community's data infrastructure. The guidelines I offer are grounded in critical approaches to data literacy that have been proposed by D'Ignazio & Bhargava (2016) and Tygel & Kirsch (2016). Second, these workshops offered a space for community leaders to articulate specific needs and values they would like to see embodied in the data dashboard. The process of eliciting these local values are a response to L. Taylor (2017) and Le Dantec et al's. (2009) call for a more local conceptualization of values when designing technologies. I

describe these values in the final section of this chapter and use them as hypotheses when reimagining the Westside community's civic data infrastructure in Chapter 6.

I begin by describing different conceptualizations of data literacy and explaining my reasons for choosing the data infrastructure literacy frame over other others. I then draw attention to specific breakdowns I encountered in this process and the consequences they had for the workshops and the dashboard's redesign. The workshops thus served as boundary objects that not only helped reveal instances of data infrastructure breakdowns, but also surfaced local values and needs that helped me rethink the Westside community's' data infrastructure. I use these breakdowns, values and needs to identify additional elements of the human infrastructure that are involved in building the community's capacity to use the data and the dashboard. In doing so, I focus on the terms *local values* and *use* in my research question, which is, *What are the local values and infrastructural arrangements that are required to build, use and maintain equitable data infrastructures that enable communities to benefit from the publicizing of data through dashboards?*

5.1 Why Data Literacy?

Data literacy has been identified as one of the main barriers that need to be addressed in enabling communities to access, interpret and use data towards their civic, personal or professional goals (Balestrini et al., 2017; J. Carroll et al., 2018; Letouzé et al., 2015; Puussaar et al., 2018). When defining data literacy, there is a tendency to highlight the technical skills that allow one to use data as a resource while ignoring the social critique of data that can lead individuals towards emancipatory end goals. Scholars have argued that data literacy includes both the computer and data science skills needed to manipulate

data, as well as the capacity to critically analyze them through a social lens (Tygel & Kirsch, 2016). It is more than just the acquisition of knowledge and skills on how to use data but needs to include the ability to read, work with, analyze and argue with data as part of a larger inquiry process (D'Ignazio & Bhargava, 2016). It includes the ability to ask and answer real world questions through an inquiry process that also acknowledges the ethical aspects of data use (Wolff et al., 2016). Within community settings, Frank et al. (2016) argue for considering a variety of data literacies that allow individuals to specialize in either the finding, manipulation or presentation of data. Extending the definition to focus on inclusion and the capabilities of the community to produce, engage with, communicate and use their own data, Letouzé et al. (2015) define data literacy as "the desire and ability to constructively engage in society through and about data".

What these scholars have in common is a belief that all approaches to data literacy can and must be leveraged to empower citizens. A data literate society, for them, is a more inclusive society; having the required data literacy skills will improve the likelihood of the community's ability to create data-based narratives that support their civic engagement needs. The aspirational goal here, which is in line with my research agenda, is to frame data literacy as a process through which the community can learn to use data to constructively engage with issues they care about.

Literacy has historically been a top down endeavor, where the elites in power set the agenda for what is to be taught and learned by the rest of society, which is counter to its spirit of enlightenment and empowerment (Letouzé et al., 2015). Paulo Freire calls for foregrounding the social injustice perpetuated by existing power structures and the role education can play in resisting it (Freire, 2000). His concept of Popular Education

advocates for establishing dialog between the teacher and the pupil to contextualize learning to the specific circumstances the pupil is familiar with. The goal here is to engage the pupil in a critical understanding of reality which might ultimately lead to emancipation.

This critical approach mirrors the reimagining of medial literacies towards *civic intentionality* that Mihailidis (2018) argues for. In acknowledging media literacy's inability to respond to the current partisanship, distrust and polarization prevalent in today's media ecosystem, Mihailidis puts forth a set of values that he thinks should help guide future media literacy initiatives. These values, which include "agentive action-taking, evoking a caring ethic, inspiring critical consciousness, developing persistent engagement, and creating conditions for emancipatory communication, where people are able to work together to respond to social problems that prevent progress from taking place" are intended to move people beyond the mere critique of media and towards civic or social impact.

Such a move towards emancipation and impact through data literacy has also received scholarly attention (Bhargava et al., 2016; D'Ignazio & Bhargava, 2016; Frank & Walker, 2016; Wolff et al., 2016). Bhargava et al. (2016) have used this approach in their effort to build *critical data literacy* with students at an alternative school in Belo Horizonte in Brazil. The concept of critical data literacy builds on Freire's model of education (Tygel & Kirsch, 2016). The focus here, is on both the technical skills required to manipulate and make sense of the data as well as the critical dexterity that ask questions and allows learners to make arguments and tell stories with data. Tygel &

Kirsch (2016) propose *data reading, data manipulation, data communication and data production* as the set of abilities one needs to use and produce data in a critical way.

In addition to the technical skills and the critical social lens, Gray et al. (2018) have argued to broaden the definition of data literacy to include the "social, historical, cultural and political settings" in which data are created and used, in order to highlight the different ways such infrastructures influence the data. They call for an infrastructural inversion (Bowker, 1994; Bowker & Star, 1999) of our data infrastructures so we can comprehend the socio-technical manner in which our data have been "cooked" (Bowker, 2013), which could ultimately lead to an infrastructural imagination (Bowker, 2014) of how they might be organized differently.

A consequence of introducing the term infrastructure is that it connects data literacy to previous scholarship from the field of STS, which has opened up the black boxes within large socio-technical systems like electric power grids (T. P. Hughes, 1984), weather systems (P. N. Edwards, 2010), health informatics (Hanseth & Lyytinen, 2016) and the like. Data infrastructure literacy, as proposed by Gray et al. (2018) therefore asks us and our learners to critically reflect on and analyze the different infrastructural black boxes that are implicated in the creation of data. This was the approach I followed when organizing data literacy workshops with the Westside community leaders.

5.2 Data Literacy Workshops as Boundary Objects

The data literacy workshops I organized served four purposes, which were of practical and scholarly relevance. First, I hoped to understand the data needs of the community leaders and how they were being met through the dashboard. Considerable amount of time had passed since work on the dashboard started in 2013, to when it was released for

use in 2016. I wanted to know if community leaders were using the dashboard and for what purposes specifically. This would help the WCA better align the dashboard to meet the needs of the community leaders. Second, I wanted to expose participants to a variety of tools and techniques they could use to make sense of data in general, and especially those available on the dashboard. This was to ensure that the dashboard could be used to its full potential. Third, the workshop activities would allow participants to perform an infrastructural inversion (Bowker & Star, 1999) of the dashboard, so they could identify and critically reflect upon the different infrastructural black boxes that make up the data dashboard. This followed from my desire to broaden the conceptualization of data literacy to data infrastructure literacy (J. Gray et al., 2018); to include not only the tools and techniques required to make sense of data, but also a critical understanding of the socio-technical, political, historical and cultural factors that impact their collection, organization and use. Fourth, and as a result of the infrastructural inversion, I hoped these workshops would offer insight into how the dashboard infrastructure could be reimagined. Such an infrastructural imagination (Bowker, 1994) would help us reimagine the future of the dashboard and also the basis for how data infrastructure literacy can be evaluated.

Within data infrastructure literacy, evidence of learning comes not from pre/post tests that measure efficient use of tools and accurate analysis of the data, but by performing a critical scrutiny of data that is reflexive, does not take data as given, and utilizes an infrastructural imagination to generate inventive uses of data (J. Gray et al., 2018). At the workshops, this came through in the kinds of questions participants asked and the qualitative nature of the conversations during the feedback sessions.

In addition to offering evidence of data infrastructure literacy, the workshop activities also offered insight into the gaps that existed in the dashboard and its underlying data infrastructure. More specifically, it was during key moments of breakdown during the activities that the scale and scope of the data infrastructure and its design issues came to the fore. This supports Star's thesis that infrastructures become visible only upon breakdown of routine functioning and mal-alignment with the concerns of its publics (S. Star, 1999). These workshop activities thus served as boundary objects for both data literacy advocates and designers of data infrastructures.

5.2.1 The Concept of Boundary Objects

The concept of boundary objects was first introduced by Star & Griesemer (1989), when they were motivated to understand how groups that belonged to different social worlds worked without consensus. Boundary objects, as they claim (i) have an interpretive flexibility, which means they reside in different social worlds with each group having their own understanding of what these objects do and how they work. (ii) maintain a common identity that both groups share, while also being tailorable to the groups' specific needs (iii) afford a back and forth, so when these groups are not co-operating or working towards specific goals, they switch between both forms of the object.

While almost anything that is interpreted differently between groups can be a boundary object, one needs to be cautious in extending this concept to workshop activities. Star has clarified that the ability to offer insights at the organizational level is key when considering what can be classified as a boundary object (S. Star, 2010). The reason for classifying the workshop activities as boundary objects is because it was

during these activities and the resulting moments of breakdown that participants critically reflected on the entire data infrastructure.

Rajapakse et al. (2018) have used boundary objects to discuss how people with disabilities build personal infrastructures to support their everyday lives. The researchers worked with individuals with disabilities and their families to co-design artifacts like video stories, photo-series, personal profiles and visit logs. These artifacts were treated as boundary objects as they served multiple groups in the project, each of which had their specific needs and goals. For the families involved, these design artifacts helped communicate and clarify the personal infrastructural needs of their disabled children with other student designers, makers and disability service organizations that were involved in the project. The co-design process helped the researchers and designers learn about the capacities, dreams, challenges and limitations of their participants and negotiate what features they would like to include in the prototype. Most importantly, this process of infrastructuring went beyond mere end-user participation as it gave the individuals with the disability more control over how the artifacts and prototypes were developed, while also enhancing their self-determination.

In another study, Wood et al. (2017) staged urban walks as a way to perform an infrastructural inversion of the Global Navigation Satellite Systems (GNSS) that all of us, as users, take for granted in our everyday use. Participants were given a brief overview of how GPS (Global Positioning System) and GNSS systems have been developed and were then asked to navigate through a collection of concrete buildings using the *GPS Test* app on an Android phone that was given to them. The app was used to defamiliarize the participants from the GPS apps they were accustomed to, as it revealed the position and

signal strength of the different satellites that were used to triangulate the position of the participants. The act of walking around the buildings by looking at a non-traditional GPS interface was a non-trivial task, which was compounded by the moments of breakdown that took place when the phone could not connect to one of more of the satellites. It was through these activities and the resulting moments of breakdown that participants critically reflected on the underlying infrastructure of GNSS and on how their own sociotechnical practices were impacted through this unfamiliar interface. The method of inverting the infrastructure also revealed alternate user experiences and offered the designers a way to critically reflect on some of the hidden issues at play. These staged activities thus served as boundary objects for both the participants and the designers.

I mention these two projects here because my approach to the data infrastructure literacy resonates with both of them. In choosing the activities for the workshops, I tried to strike a balance between offering the participants data about the community that they were familiar with versus that which they knew nothing about. Not all of these activities went as planned and there were instances in which the activities broke down. Some of these were caused due to *irrelevant data* and others due to *inaccurate data.* It was during moments of breakdown that the data infrastructure lent itself to the most critical reflection by the participants. This included the socio-technical, historical, cultural and political settings within which data were collected, cleaned, curated, controlled, combined, contextualized and conveyed. These moments of breakdown presented me with design provocations I used to reimagine the data infrastructure for the community. In the next section, I focus on these moments of breakdown that I experienced during the workshops. I then reflect on what this means for how community leaders make sense of

data and what we as design researchers can do when supporting community data literacy efforts. The table below offers pseudonyms and affiliations of the community leaders who participated in the workshops. I have chosen not to include details of four participants who attended only one workshop as I have not used any of their data in this dissertation.

Table 3: Workshop participants identified by their pseudonyms and organizational affiliations

Name	Organization	Workshops Attended
Stephanie	SWEEAC (The Southwest Ecumenical Emergency Assistance Center)	1,2,3,4
Rhonda	SWEEAC (The Southwest Ecumenical Emergency Assistance Center)	1,2,3,4
Bluma	TryCope (Try Reaching Youth Through a Community of Parenting Effectiveness)	1,2,3,4
Carol	TryCope (Try Reaching Youth Through a Community of Parenting Effectiveness)	1,2
Kimberly	Sisters Action Team	1,2,3,4
Dayana	Sisters Action Team	1,2,3,4
Isabel	WAWA (West Atlanta Watershed Alliance)	1,2,3,4
Stella	Historic Westside Heights/ Bankhead Neighborhood	1,2,4

	Association Inc.	
Hailey	Historic Westside News	5,6,7,8
Patricia	The Conservancy at Historic Washington Park	5,6,7,8
Clara	The Conservancy at Historic Washington Park	5,6,7,8
Casey	Historic Westside Gardens	5,6,7,8
David	North West Ministerial Alliance	5,6,7,8
Janet	North West Youth Power Inc.	5,6,7,8
Mary	Raising Expectations	5,6,7
Roy	NPU T	5,6,7,8
Kiara	WAWA (West Atlanta Watershed Alliance)	5,6,9

5.3 Moments of Breakdown

5.3.1 Irrelevant Data

The first warm up activity at the workshop was to build data sculptures (D'Ignazio & Bhargava, n.d.-b). To situate the participants and give them a sense of what they could make, I played the video that described the activity from the data culture website. For materials, we had foam blocks, pipe cleaners, colored pom-poms, colored plastic cups, googly eyes, glue sticks and construction paper. To build their data sculptures, I asked the participants to pick one of three example datasets that come with the activity: most popular ice-creams in the US, how does Somerville feel or the cost of education at MIT.

I intentionally chose datasets that had no association with the Westside because I did not want participants to get absorbed in discussing the meaning, implications and issues with the data for this first warm up activity. In the first two exploratory workshops that I did in 2017, I used business data about the Westside within all of the planned activities. This resulted in conversations about the quality of the data, its visibility, impact and other issues that took time away from learning about the tools and techniques to analyze data with. While these conversations were insightful, participants did not learn anything new about *data manipulation* and *data production*, which are two of the four aspects of critical data literacy (Tygel & Kirsch, 2016). The workshops were structured to have a mixture of activities where some were about learning to use tools to manipulate data and produce interesting outputs, while others were about performing an infrastructural inversion of the data. This first activity was meant to be a short icebreaker where participants thought creatively about creating physical representations of data with craft materials.

Figure 18: Building data sculptures with craft materials.

All participants tried to work with the data from the given examples to create their data sculptures. Casey though, had trouble relating to the data in the handouts. She ignored the data in the handouts and built a data sculpture that highlighted an issue that was of concern to her, that of rat infestation in her neighborhood. The colored blocks in Figure 19 represent the different houses, some of which are vacant and some occupied. The vacant ones can be identified by the overgrowth, which is represented with the green pipes. The big googly eyes are people and the small ones are rats. *"Some of the buildings*

have rats and people. Those eyes are bigger because they're probably high!" she joked. The black dots are the dogs and the small dots are all the germs.

Figure 19: Data sculpture of rat infestation in the neighborhood.

This issue about data in the workshop activities not being relevant to the participants came up again during the 'WTFcsv activity'. With this activity again, my goal was to draw attention to the different online tools that can be used to quickly generate preliminary insights about data. This was again the reason I chose to use the default datasets presented on the databasic.io website instead of using datasets describing the Westside. Mary and her group did not seem very engaged when working on this activity. Although they did follow the instructions that were given, they also seemed frustrated that they were asked to look at fictitious data that had no relationship to the Westside.

They voiced their frustration even in the activity handout that was given to them (Figure 20).

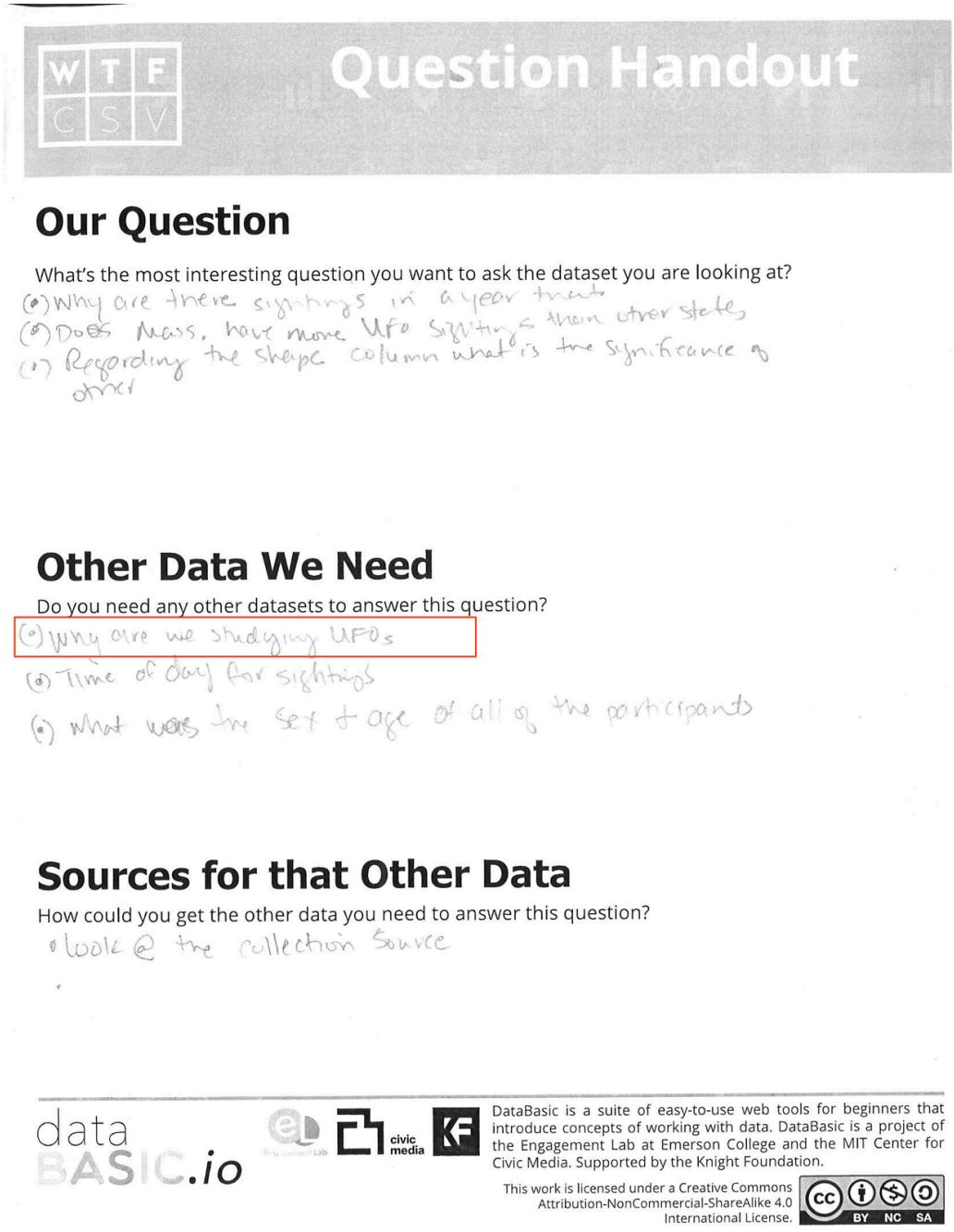


Figure 20: Handout for the WTFcsv activity.

5.3.2 Inaccuracies and errors in the data

Before starting the 'Information Hunt' activity with the dashboard, I did a quick demonstration of all the data that were available in the different modules. When presenting the Public Safety module, a visual about the number of burnt structures seemed to stand out to the group. The visualization showed that there were eight burnt structures in the Westside neighborhoods, which was far fewer than what the participants knew it to be. Participants who lived these neighborhoods knew the staggering rate at which old and abandoned homes were being burnt down over the past few years. To see these numbers being under reported on the dashboard made them question not only the usefulness of the dashboard as a resource, but also the values of organizations like the Atlanta Police Department (APD) which provided the data.

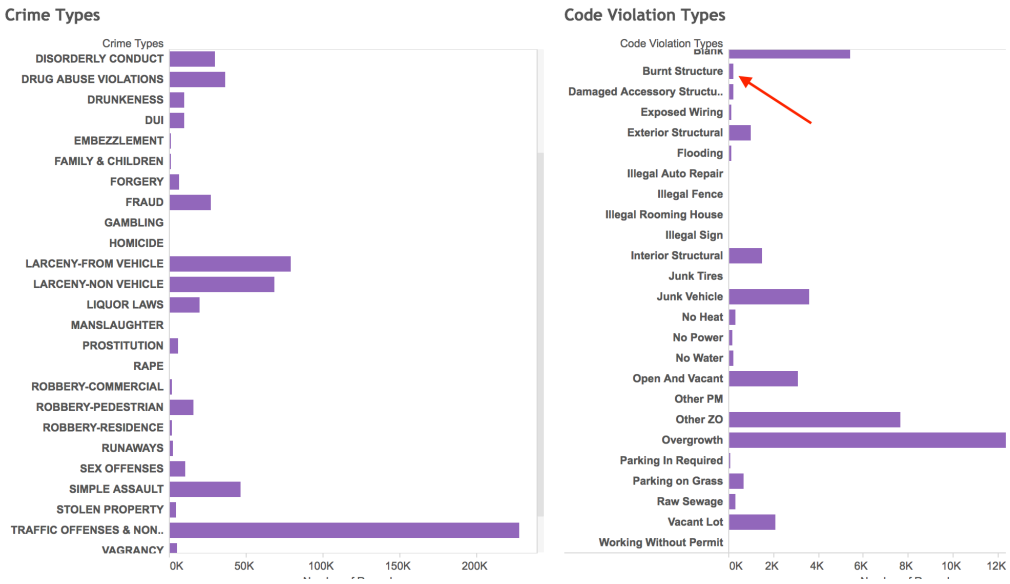


Figure 21: Visual from the Public Safety module showing an inaccurate number of burnt structures.

The number eight represented an oversimplification of the phenomena where, the community leaders insisted, owners were intentionally burning down their properties to

claim insurance on them. We spent some time trying to decipher what the number represented. Was it the number of burnt structures for a specific NPU for a specific year or did it represent all of the burnt structures in the Westside? Given that the data came from the Atlanta Police Department (APD), some asked if this was an error in the way the data was collected and classified by the APD, or in the way the data analysts at Georgia Tech had understood it or merely a technical glitch in the interface? Some even went as far to claim that the APD had intentionally manipulated the numbers in order to highlight an upward trend in the community's public safety.

This was a strong claim and demonstrated the adversarial relationship that the community shared with the police department. The residents not trusting the data being put out by the APD was one of the reasons they asked for a Public Safety module that had data they could use to hold the APD accountable. While there is no evidence of this data being manipulated by the APD, it is worth dwelling upon the reason behind this error and the participants' reaction to it. The relationship that the Westside shared with the APD and the context that the data claimed to be representing led to an infrastructural inversion in which the community leaders unpacked these inaccuracies and treated them as local signifiers of data context and provenance (Loukissas, 2016). Data are always connected to place (Crivellaro et al., 2016; A. S. Taylor et al., 2015), and this criticality regarding data provenance where participants not take data as given is a desired quality within the mandate for data infrastructure literacy (J. Gray et al., 2018).

The inaccuracy of the data on the dashboard also made Casey question the trustworthiness of the tool itself. Patricia and Roy disagreed and were of the opinion that even such inaccurate data was empowering, as it gave them the evidence they needed to

go to the city department (which was the source for the data) and ask them why the numbers were what they were. Hailey agreed and offered her perspective, in which she believed city officials in Atlanta were using an inaccurate Area Median Income (AMI) numbers when calculating the rent and number of affordable housing units in the Westside. The median income for the Westside neighborhoods is in the $400-$500 per month range, which is 25-30% of the AMI. But city officials were classifying units as affordable if they were within the reach of those who made 60% of the AMI. This error was because officials were including incomes from more affluent metropolitan areas of Atlanta, where the average income was much higher. Learning about this discrepancy allowed community leaders in the Westside to approach the city officials and request changes to be made to the manner in which affordable units are stipulated. Thus, Hailey argued, inaccurate data could also be empowering, but this was still not enough for Casey to trust that the dashboard as a useful resource.

This example represents contrasting points of views through which both Hailey and Casey displayed one aspect of what constitutes critical data literacy as defined by Tygel & Kirsch (2016), viz., *data reading*. Data Reading involves recognizing that data is not an objective fact, but the result of a sociotechnical negotiation by its creators. While Hailey saw inaccurate data from the glass half full/half empty perspective, Casey was unwilling to put her trust in something that she believed was inaccurate. Hailey believed it was part of the community's collective responsibility to question the data and hold its creators accountable. For her, inaccurate data was just as valuable as accurate data, as it allowed them to dig further into the process through which it was created and understand the reasons for its inaccuracy. For Casey, this inaccuracy was the tip of the iceberg that

made her question the utility of the dashboard entirely. If this was inaccurate, it was likely that there are other data on the dashboard that are also inaccurate. Both of these perspectives ask critical questions of the data and ask its creators to be accountable for the kinds of data they release. Casey's trust in the dashboard and the team behind it was further eroded at the next workshop during the 'Remix a Visualization' activity.

The 'Remix a Visualization' activity (Bhargava, n.d.-b) asked the participants to remix a given visualization to meet the needs of a specific audience they had in mind. Because participants at the previous workshop were unwilling to engage with data that was not relevant to them, I made sure that the data for this activity was relevant to ongoing issues on the Westside. The visualization I presented was about funding that different organizations on the Westside had received from The Arthur M. Blank Family Foundation, which since 2013 has committed a total of $37M towards positive transformative efforts taking place in the Westside neighborhoods [source: Blank Foundation website[16]]. The visual was created by a student at Georgia Tech as part of a project studio.

[16] https://blankfoundation.org/investments/

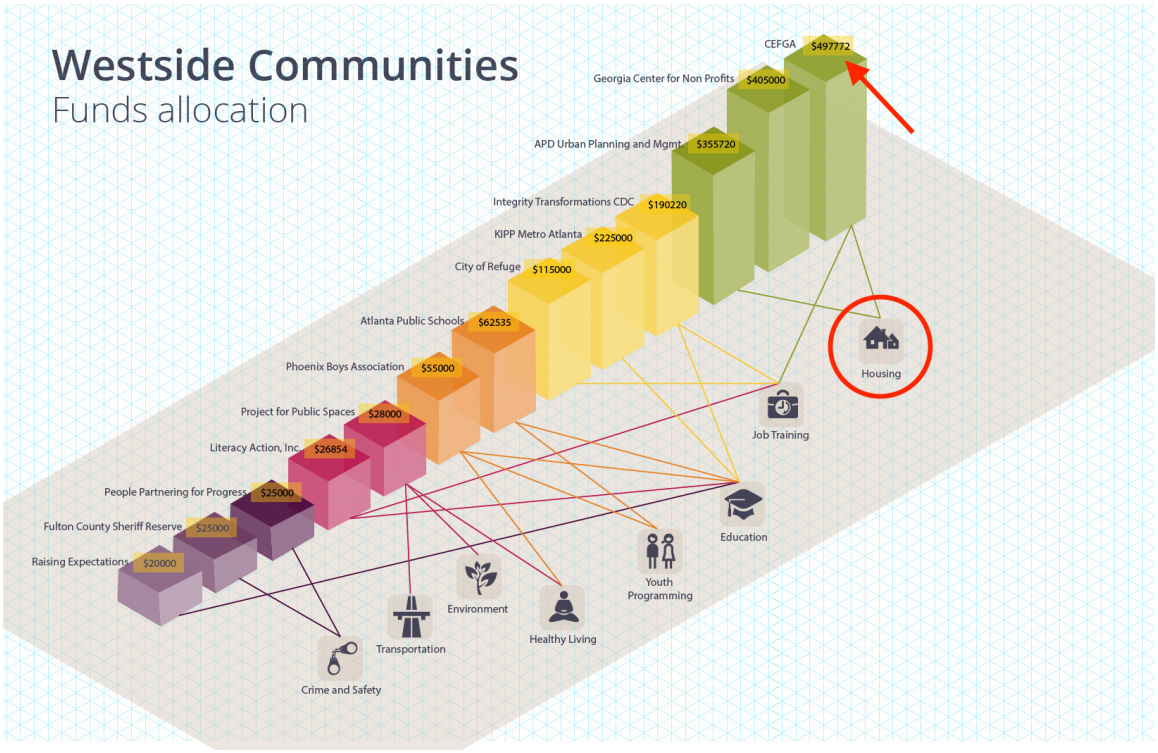


Figure 22: Westside Communities Funds Allocation Visualization. Done by a student in the Public Design Workshop at Georgia Tech.

Looking at the numbers in this visual got everyone in the room really upset about how data was being used to tell a story that was completely different from the ground reality they had been experiencing for years. The point of contention was that the visual showed close to a million dollars given to two organizations, (Construction Education Foundation of Georgia (CEFGA) and APD Urban Planning and Management LLC, which were classified as working on "housing". This number would seem admirable to anyone outside the community but infuriated those in the room who were still helping out homeless teenagers, adults and families on a regular basis. They could not fathom the fact that close to $1M had been spent on housing since 2015, but the problem of homelessness remained unabated. Hailey, put it across most succinctly in Figure 23 (left), which showed how all this money had zero impact on the neighborhood. Figure 23 (right) by

Janet asks for need, impact and results of all the money that has been invested into these organization between 2015 and 2017.

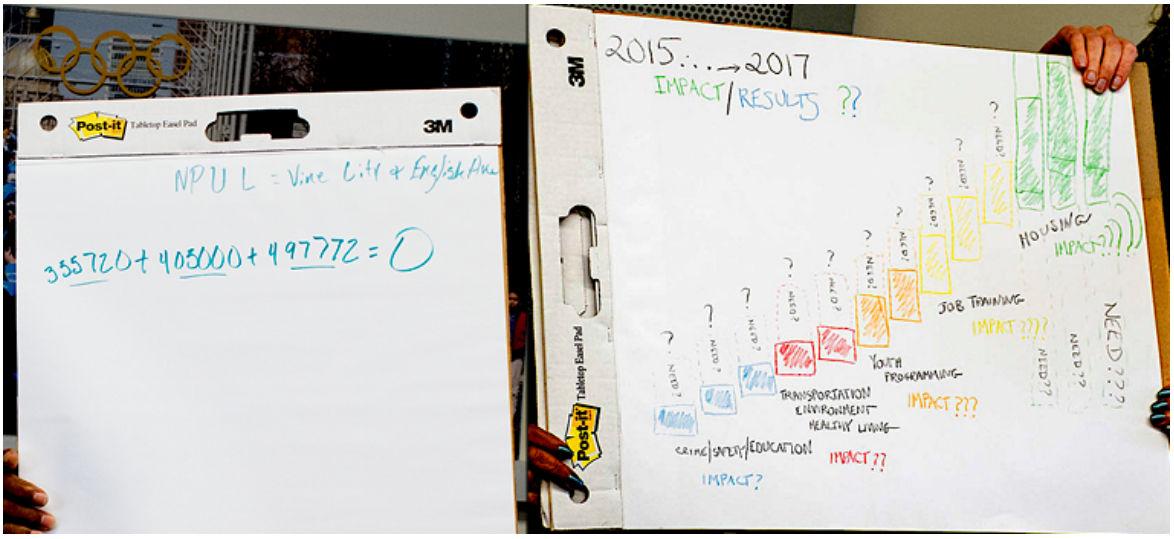


Figure 23: Remixed Visualizations ask about impact of funds given to different organizations.

CEFGA, which received $497,772 is classified under both ‘housing’ and ‘training’. But CEFGA’s website states that it is a 501(c)(3) non-profit that specializes in offering construction training to high-school and adult students. The organization offers job-training and placement services, which are better classified under ‘training’ rather than ‘housing’. This error in classification led to quite an animated discussion among the participants, with many of them being visibly upset about the alternate reality that this visual was putting forth. Participants at the workshop did not know what CEFGA was. When Hailey asked the group what the acronym stood for, Casey, who was frustrated at what she was seeing in the visualization replied “B.U.L.L.S.H.I.T”!

Classifying CEFGA as ‘housing’ was likely a mistake on the part of the student. The student probably didn’t read too much into the $497,772 that were allocated to CEFGA. The workshop participants were more critical of this number and demanded to know why

that money had been allocated to that organization and what they had done with it. The questions of accountability and impact that they were asking through their remixed visualizations are examples of *data communication*, which is another facet of Tygel & Kirsch's (2016) definition of critical data literacy. Data Communication refers to the ability to determine the best format in which one might be able to communicate data in a responsible and precise way, while avoiding any misunderstandings. Determining what to communicate required a level of contextual familiarity that the student lacked. Knowledge of the context that the data was representing led the community leaders to inventive data practices, which is also an instance of data infrastructure literacy (J. Gray et al., 2018). Other instances of data infrastructure literacy came forth through the questions and discussions that took place as a consequence of each of the breakdowns described so far.

5.3.3 Consequences of the Breakdowns

Discussions during the workshops revealed that while there was a strong commitment to using data for advocacy, not everyone was equally skilled in doing so. As participants at the workshops, these community leaders were wearing several hats at the same time. Participants were at times representing their organization, their specific geographic community, be it their street, neighborhood, zip code, city and even their race. As the rest of this section will illustrate, participants were cognizant of the issues in each of these capacities and had ideas about the data they needed to advocate for change but did not always have the means to collect them, analyze them and produce narratives in a timely manner. They also realized they needed more than just personal stories if they needed to affect policy level changes.

The groups willingness to participate in the data activities seemed to have reached its limit by the final workshop. As a result of all the breakdowns that took place, Hailey was particularly disheartened to see how data could be twisted to tell different stories. At the start of the first workshop, she had stated that she was interested in working with data because *"data don't lie"*. By the end of the fourth workshop, Hailey believed her lack of engagement with data was allowing others to portray the neighborhood differently from how the residents were experiencing it. As a representative of the Westside's monthly newsletter, she felt she hadn't done enough to convey the community's story through data. Similarly, Casey feel personally responsible for not being able to use her skills in video production to convey her community's concerns and issues to the outside world.

While it was my intention to structure the workshop activities in a manner that offered a rich context within which to reflect on the data infrastructure, it did lead to some participants feeling manipulated and frustrated. Hailey mentioned she had signed up for the workshops to learn how to use data, but what she experienced was kind of manipulation similar to what their community had been experiencing with for years. I believe she might have been referring to the history of development projects and policies that were designed to benefit the wealthy and powerful at the expense of the Westside residents. Casey was frustrated that she was sitting in a room doing "coloring book stuff" while there were people out on the street who were homeless and whose life was at stake. The only difference she mentioned, was that this time *"we're being paid to be frustrated"*.

This issue of homelessness was given a face when Sally, who was once a homeowner revealed that she had been homeless for the past week and was relying friends to support

herself. Mary, who accompanied Sally said they had been looking for a place for the past three weeks. Casey learnt about this and was trying to find Sally a place to stay. Hailey and others consistently offered their homes to kids in the neighborhood who didn't have a place to sleep at night. This was part of the reason they had a strong reaction to the Blank Foundation visual which showed close to one million dollars being spent on housing. The visual glorified the philanthropy of the foundation and said nothing of the impact these funds were having on the residents.

Another consequence of all of these breakdowns was that by the time we reached the fourth and final workshop participants seemed tired and reluctant to continue with the activities. Casey wanted to instead talk about what I had learned from the workshops so far, and how that would impact them as a community or the design of the dashboard going forward. She asked in a tone that was firm and reflective,

> *"We're talking about our experiences personally as individuals and then our relationship with the community. So then I want to know from you hosting these workshops, what are you taking away from this experience."*

I was a little taken aback by the question and didn't know exactly how to respond. My initial response about my motivation to engage with the community through their matters of concern seemed unsatisfactory. To clarify, Casey went on to describe how she had lost sleep over the Blank Foundation visual I had shown at the previous workshop. She wanted to know why I had shown them a visual that *"derailed the class"*. She seemed to think that I had not studied the data or the visual enough before showing it to them at the workshop, which she believed was irresponsible. She believed that their collective reaction to the visual made for great television rather than data literacy and wanted to

know if that was the reaction I was expecting to when I showed the visual. She wanted to know if I had learned anything from that experience that would inform my future work. When others asked Casey to clarify what exactly she meant, she pressed on.

> *"As a person that's putting out the data, now, after you've spent this time with us, will this shape the way you put out data?"*

While I stood there trying to process my thoughts and respond, Janet clarified Casey's question further for me

> *"What for you personally has shifted for you [sic] or has moved for you? Or how much you factor in now that…(you) know that this dashboard is going to be a repository for data that gets to be put out and as you see our reactions to some existing data or datasets and kinda our real angst about what currently exists as data."*

Now that I reflect on this exchange, the question seems straightforward. The participants had every right to know how the insights I had gathered from them would be reflected in my next steps. These community leaders had taken time away from their overbooked schedules to participate in the workshop sessions, so it was important for them that their expertise be acknowledged and represented in my next steps. They didn't expect me to be personally accountable to the community as they were, but they did expect me to at least be accountable for the data I was putting out. Once I had some time to process their question and understand where they were coming from, I responded saying

> *"The thing that's really changed for me is understanding how my view, in spite of being involved in this project for so long, is not the same as the perspective that*

you guys can offer. The example of the visual we saw last week. I mean, I looked at it, I looked at the breakup of funds, I saw that, I recognized some of the organizations and some of the work that's happening, but the stories that you guys got out of it were completely different, right. You know what's happening. You were talking about impact, what work has happened, why aren't we seeing results. So those kinds of conversations are really important and that's something that I really appreciated throughout this process of engaging with all of you, understanding what the real issues are, and the reasons behind that. That's something that can only come through this kind of engagement and that's something I really appreciated and something that I really want to do more of going forward."

Was this response sufficient? I could not tell for sure, but Casey brought it up again and asked me again towards the end of the session, what my responsibility to the data was. Was it responsible for me, as someone representing the dashboard to put out data that was wrong and misleading and ask the community to engage with them? And what would I do now that I knew about the erroneous data?

Answering these questions required me to be reflexive about my role in creating these data experiences. While my intentions were to encourage discussion and bring about a greater awareness of the infrastructural journey data go through, I had not considered the violence these efforts might cause to the communities for whom partial perspectives and manipulation part of their lived experience. The data I was presenting about crime, burnt structures, and housing expenditures put forth a reality that these community leaders were fighting to change. My intention to use such data to further their data infrastructure

literacy seemed manipulative and a violation of their trust. What would it take for me to do justice to the data, to the variety of stories that come out of it, and the way in which the community leaders experienced the data personally? I will take up these questions when I address issues of data justice in section 5.5. For now, I return to the workshops and describe how they helped me meet the other goals I had in mind, which were to identify if and how the dashboard was meeting the community's data needs.

5.4 Eliciting Community Data Needs

Participants at the workshops expressed their data needs through the activities and in the follow-up interviews that I did with them. I interviewed 6 of the 7 participants who attended the first round of the workshops (in Feb/March 2018) to learn more about their data needs. During the second round of workshops (June/July 2018), I included the following questions about data needs as part of the workshop activities.

Q1. What is the problem/issue you are dealing with?

Q2. If you are using data, how are you currently using it to solve your problem? Can you list your datasets?

Q3. Do you have all the data you need?

Q4. What other data will help you tell your story better? Does this data exist? Where can we find it?

The participants spent some time discussing these questions and then wrote down their responses. For Hailey who was involved in running a community newspaper, this translated to *"something as simple as a contact list"* or *"something like the number of people in the community that use Facebook to interact*", which they didn't have. She could use this data to analyze the paper's current readership and plan effective strategies

to expand it. Patricia was interested in knowing how the neighborhood recreation facilities, which included a playground, a baseball field, tennis courts and a natatorium, were being used and the demographic breakup of its users. This included data about the number of seniors who lived in the neighborhood, the number of homeowners vs rental residents and the number of homes that were on the market to be sold. This stemmed from her concern about keeping the facilities relevant to a neighborhood that was rapidly changing due to the ongoing real estate development projects. David's organization had worked with students from Emory University to collect data about how many after school programs were running in their neighborhoods and how many of them served the youth (pre-k, Head Start, elementary). His organization collected this data so they could identify gaps and request funding for more such afterschool programs. The data had not yet been used though, because as a small organization, they lacked the time, capacity and expertise to analyze the data, identify specific funding sources and present their insights in an appropriate format and within the necessary timeframe. Roy needed data about how many residents were engaging with their NPU's (Neighborhood Planning Unit) to justify to the mayor or the city council that they should triple the NPU operating budget that year. There were other data needs that came up during the conversation and during the interviews I did with the participants. I have listed them here in table below to offer an overview of all the requested community data needs.

Table 4: The Westside community's data needs.

Issue	Data Used/Needed	Is Data Available?
Affordable housing in the Westside neighborhoods.	Housing affordability data and visuals from the	Yes. Data comes from the American Community

	Westside Future Fund Data Dashboard	Survey.
Students' performance in elementary schools	Tracking students' performance through elementary grades in school.	No
Orientation booklet for churches.	Church collects and maintains data about their programs. Need more data about families in need.	Yes, but needs to be expanded to include more neighborhoods and families.
How do culture and arts impact youth and pride in the community?	Need to collect more data about impact	Not sure where to find it
Where can we put community gardens in the neighborhood?	Need data about homeowners. Are they planning to sell? Would they like a community garden on their property?	No
How many youth service programs are running in the neighborhood?	Some data collected by Emory University students.	Data collected but not analyzed or used.
Data to help the community newspaper reach a broader audience.	List of subscribers. How many of them use Facebook?	Data not collected.
How are a neighborhood park and recreation facility being used?	Using attendance numbers from camps and usage of natatorium. Need more data about neighborhood demographics like number of seniors, homeowners, renters, homes for sale etc.	Data available for American Community Survey and the dashboard but they are not the most current.
Data about resident	Health, income and	Data available via Fulton

engagement with NPU's.	demographics data. Need help converting data into insights and apply for funding.	County, Georgia and Census websites but they are not the most current.

The process of learning about the community's data needs also revealed how the dashboard was meeting them. It was clear that the dashboard was fulfilling a need when it came to finding data related to demographics and schools that came from the census. As David said after I completed a demo of the dashboard,

> *"For me, in the faith-based community we really really need this kind of information. As one pastor would say, going on feelings just does not work. You need information."*

What David implied is that one cannot advocate for change by relying only on impassionate pleas. These pleas need to be backed with data, the kind that was available on the dashboard. But the dashboard was also limited in what it could offer. Community leaders needed data that was current and up to date. At the time of the workshop in 2018, the dashboard had data from 2015. Planned updates to the data could not be completed due to the organizational changes that the dashboard was going through. In other instances, the data these leaders needed had not yet been collected as the issues they were dealing with were specific to their neighborhoods or organizations and outside the scope of city, county or state departments. These departments could not be tasked to find data to address concerns like planting more community gardens, about utilization of the neighborhood recreation facilities, or the impact of culture and arts on youth pride in the community. This data was hyperlocal (J. Carroll, Hoffman, et al., 2015) and collecting

them was the responsibility of these community leaders who did not always have the capacity to do so. David confirmed this when he said,

> *"We had the information, but as a small organization, number one, how do I take that information and then approach a funder with the information. My organization did not have the capacity to do that. And also, number two, if there was Fulton county funding available, they have a specific timeline that you have to acquire and ask for that funding. Same with other foundations. So there was just not enough information from my organization to take that information and go to the funder. So I would need the capacity of you [sic] to help me get the capacity to go to funders, to access funders."*

One reason for this lack of capacity could be attributed to the size of these organizations. Second, is the lack of data literacy, as described previously. Third, was the fact the fact that these community leaders just did not have the time to deeply engage with the data. The Westside communities have been under the spotlight for the last several years and are going through rapid change. Community leaders are part of several planning committees and meetings to make sure that their interests are represented during these times of change. So much so that they hardly have time to catch up on what is happening in other parts of the neighborhood. As Patricia noted,

> *"I haven't seen her 2 years, since 2015. I haven't seen you since before 2015. I haven't seen him since 2016 you know. So, that what all this is. It doesn't matter about surprises. Going forward, some of us, we make a decision, you know, find time to just give a heads up. David, it may be, from my understanding, coz I don't deal over in that area with the housing or too much of the schools at this point.*

But I did hear that our supposed housing, HUD secretary had mandated that rents were going to be tripled. And I'm assuming that that is coming down from, trickling down even more on top of everything else from so called federal government. So, I understand what you're saying. You're in something. Coz none of us can know it all, all the time. We don't have a hundred eyes in our backs."

Mary acknowledged that this lack of communication is an issue, but it is a reality she had come to accept as part of her job.

"part of that is that I keep my head down, so I'm constantly working, constantly working. And so as a consequence, sometimes I'm guilty and I don't raise my head up high enough to say I need to reach out to community partners so that we can leverage each other's strengths. I own that part. But I'll also say, that part of it is also that people who are working in the community, we have to, we have to, I force myself to keep my ear to the ground. Coz I need to know about problems before they become a problem."

This inability to stay in touch with each other and keep up with what is happening in different parts of the neighborhood also meant that the community leaders were wary about spending time attending more workshops to learn new skills. Part of the reason many of them persisted with my workshops was because it came with a $400 stipend for their organizations, which was not an insignificant amount. But attending a workshop just to learn more about how to use a dashboard seemed like a waste of their time. As Hailey said when asked about her experience of finding data on the dashboard,

"I don't want to attend another workshop to learn this. I just want it to be easy to use. I want to be able to operate it like my preschoolers get on YouTube. That's

how easy I want it to be. I don't want to take another class or workshop coz everybody's (inaudible). I just want it to be easy. That's the best way to make it accessible. I want no more of my Saturdays, no more of my nights."

Others, after going through the information hunt activity on the dashboard, also made specific suggestions for features they would like to see on the dashboard. Roy explicitly asked for the dashboard to include information about where and how their efforts could be funded by various sources.

"One thing I would like to see on the website is the monetization of the initiatives. Where the funds are, and the path to get to those funds. Coz one thing that we all share in common was a lack of capacity, which of course would be all of this put together. To input data, so we can use the Word Counter or things like that etc etc. But where are the funds and where do we get the funds from and so, if CWK could be the one stop shop, where funding availability were included in that data, that would be helpful to me."

Some saw the dashboard itself as an accountability tool, a tool they could use to hold me and others in the community accountable for the data they were putting out and the impact they were having. As I elaborated in the previous section, the community leaders held me accountable for the data I was putting out on the dashboard. But Hailey also saw the dashboard as a way for her to hold others in the community accountable for the work they were doing

"So, I'm looking at the platform as not just a way for us to collect data but for us to say, 'hey, this is wrong'. Not because I'm asking you for anything, but because

I'm holding you accountable. And I wanna tell the world. It's an accountability tool."

In addition to holding me accountable for the data I was putting out, Roy believed that it was also their responsibility to ensure that the data on the dashboard was accurate.

"When we see data and when we see numbers from our perspective, whether we agree with it or disagree with it, it's our responsibility to check that source."

Casey, on the other hand, believed that GT had more of a responsibility to ensure that the data they put out was accurate. As a reputable public institution, any data put out by GT would be taken as "*gospel*" by outsiders and media outlets who had no way of verifying its accuracy. She insisted that the community needed to be able to access accurate data about themselves in a usable format, so they can create narratives that highlight their point of view rather than those of outside researchers or media outlets.

Community leaders had come to trust the data on the dashboard thanks mainly to the public engagement efforts of the WCA over the years. In the absence of the WCA and their engagement efforts, there was uncertainty about who represented the dashboard and who at GT was responsible for maintaining it. Laura felt that the dashboard makes the data "lose ownership". When she saw erroneous data about burnt structures on the dashboard, she wasn't sure if she should contact GT or the Atlanta Police Department to get the issue resolved. Was there somebody they could talk to in order to report issues, request features or even just understand the data better?

The issues mentioned above point to more than just the data needs of the community leaders. The vignettes reveal how the community leaders appreciated the idea of being data literate and did not need to be convinced about the value of being data driven. While

there were a few who attended the workshops only to raise money for their organizations, others came there to learn more about the kinds of data and tools they can use to create their own data stories. The workshops brought community leaders together to discuss important matters of concern and talk about how they were using data to tackle them. The two contributions I make in this chapter are a result of these discussions. First, I use the insights from these workshops to propose a set of guidelines other researchers or practitioners can use when organizing similar data literacy workshops with minoritized communities. These guidelines, which I discuss in section 5.6, are meant to be a starting point for organizing workshops and will need to be tailored to the specific needs and circumstances of the communities one is working with. Second, I use the discussions to compile a localized set of data justice values that serve as hypotheses in the subsequent reimagination of the Westside community's data infrastructure. I discuss these values and situate them within the data and design justice scholarship next.

5.5 Discovering Local Values of Data Justice

Design and values are storied (Parvin et al, publishing as JafariNaimi et al., 2015; Parvin, 2018) in that it is stories that help us negotiate and identify values so we can design appropriate interventions. The different breakdowns I have narrated in previously helped draw out the nature of the situation in the Westside and the action it demands. I distil those breakdowns into a set of values, which, although at times in conflict with each other, serve as hypotheses when reimagining the Westside's data infrastructure in the next chapter. The process of identifying these values involved wrestling with a much larger set of themes, extracting common values from them and then distilling them into a manageable set of values. I arrived at this set by exercising my own 'appreciative

judgement of values', which is a term Iversen et al., (2012) use to describe the process through which designers refine values and embody them in design practice.

As described in Chapter 2, there are several frameworks and theories describing what it means to do justice to data. Equity in the distribution of data is one of the values of data justice that the WCA dashboard hoped to achieve from the start. But what does equity mean when it comes to data? And more specifically, what does data equity mean in the context of the Westside neighborhood and the data dashboard? The breakdowns and the resulting discussions allowed the community leaders to articulate several values that were important to them. For the community leaders, a more equitable data infrastructure would

1. Support the community's data infrastructure literacy.
2. Empower the community towards action.
3. Foster accountability and responsibility in and through the data.
4. Disclose the context within which data are created.

5.5.1 Support the community's data infrastructure literacy

A lack of data literacy is one of the most often cited concerns when it comes to engaging with community data infrastructures (Balestrini et al., 2017; J. Carroll et al., 2018; Letouzé et al., 2015; Puussaar et al., 2018). A Senian conceptualization of data justice also asks that we consider not only what people are capable of doing, but what they actually do with the data that they have access to (Britz et al., 2013).

Workshops are just one of the many different strategies that civic data scholars have proposed to get communities involved in creating, accessing, analyzing and using their own data (Baack, 2015; Bates et al., 2016; Puussaar et al., 2018; A. S. Taylor et al., 2015). The data literacy workshops I organized required a significant investment of time

and effort by the community leaders, which not all of them were willing to put in. While such formal learning engagements do have a place, there is also a need for more informal learning opportunities where community leaders can learn at their own pace or within their own communities of practice. Data ambassador programs, which have been suggested before by Carroll et al., (2018) is one such informal learning avenue that I explored when reimagining the Westside data infrastructure. I say more about my approach to organizing a data ambassadors program in the next chapter, and also about how its goals were at times in conflict with other values of justice I outline here.

Data literacy was desirable to the community leaders because it allowed them to analyze data on their own and create narratives that would forward their perspective on important issues. These perspectives could serve as counters to other narratives that tend to circulate in popular media. But, critical and feminist scholars of data argue that all perspectives are partial and there is no objective truth to any situation. Speaking in the context of design, they claim that all knowledge is partial, situated and located within the political, social, cultural and material context within which it is produced (Haraway, 1988). L. Suchman (2002, p. 6) argues that "The only possible route to objectivity on this view is through collective knowledge of the specific locations of our respective visions". Data literacy that can lead to such a plurality of perspectives helps expose multiple facets and interpretations of the data.

During the workshop, participants were upset when their experience of ground reality did not match up with what they saw in other interpretations of data. They felt that more powerful voices of funding agencies and media outlets were being privileged over their own experience with homelessness, which was that the millions of dollars being spent on

housing reforms had so far, had zero impact on the community. These community leaders hoped for a way they could talk back to the data in a way that included their personal experiences. But not all of them were equally skilled in all aspects of producing such data narratives. Doing justice to data would involve offering these community leaders a platform through which they could collaboratively forward their perspectives on data, one that would be open to a plurality of interpretations of the data and its stories. A data platform/infrastructure that would support the ability to create alternate interpretations of data or annotate existing ones might be one way forward.

Supporting the community's data infrastructure literacy would imply that in addition to having the data literacy and the platform to produce such counter narratives implies, community leaders would also have access to the data and know how/why they have been created. Knowing how data have been processed and with what purposes allows one to reflect on what stories the data can and cannot support. Such questions about data provenance also came up during the workshops.

Some community leaders wanted to know why the data and the dashboard was hosted on a Georgia Tech domain if it was a community resource. Could we consider more community led and controlled (Costanza-Chock, 2018) forms of data infrastructures/platforms that could convey the collective ownership of data? Was it sufficient that the data was hosted on a domain controlled by a non-profit, or should we consider other forms of ownership, like hosting them with the public library? Would such ownership imply that the community could equally partake in the benefits and burdens brought about by such data infrastructures? Benefits, as Dombrowski et al. (2016, p. 662) explain, are the "wealth, goods, opportunities, and access to resources" that are brought

about by such sociotechnical systems, while burdens include the "undesirable work, taxes, lack of adequate income, and environmental pollution" and other negative impacts. Such ownership is more likely to combat the injustices perpetuated by data, as it gives the community access to, and the ability to influence the underlying data, structures, databases, and processes and politics that govern them (J. A. Johnson, 2016). Included in this ability to share in the benefits and burdens of data is the autonomy to choose if and how the subjects of these data are to be included in these technologies (L. Taylor, 2017).

5.5.2 Empower the community towards action

Empowering the Westside community with data is one of the main goals the WCA hoped to achieve through the dashboard. While the dashboard does a good job of bringing a variety of datasets under a common interface, it is up to the user to ask relevant questions, interpret the data and present them in suitable ways. Such a "view from nowhere" (Haraway, 1988) that merely visualizes all the data but does not answer any specific questions might be disempowering communities rather than empowering them. Empowering data stories must not only help reveal previously invisible perspectives, but must also lead the community towards action (D'Ignazio & Klein, 2020; Dörk et al., 2013). It requires that we do more than just relate the data *as is* (Parvin, 2018), but offer customized snapshots of data that could lead to specific actions for the community, rather than the individual user (D'Ignazio & Klein, 2016). In addition to facilitating action, empowerment also allows communities to choose the specific actions they would like to enable through data. One must ask if a dashboard even the right tool to consider when the goal is to empower communities where data literacy skills are not equally distributed?

This question becomes all the more important when one considers the amount of work that goes into not only building, but also maintaining the dashboard with updated data. Maintaining data is part of the invisible work that goes into keeping a data infrastructure usable and is seldom accounted for (Crain et al., 2016; M. Muller et al., 2019). In the case of the dashboard, development, improvement and maintenance efforts were put on hold whenever funding was hard to come by. In extreme cases, like during the organizational reshuffle of the WCA, the dashboard was taken offline altogether. If the dashboard is to serve as a resource for data equity, it's availability cannot be subject to institutional power struggles and academic funding cycles. The dashboard and its underlying infrastructure need to be consistently available and maintained for continued use. Structuring the dashboard so it relies more on the community's distributed data infrastructure would make it less reliant on the vagaries of a single institution. Such an infrastructure would include an alignment of the humans, materials, technologies and organizations that can support the collection, curation, cleaning, control, collaboration and creation of data and its stories.

Empowerment through data also needs to contend with what happens when those who actively build and maintain the data are not available to perform their duties anymore. Sustainability is a core concern within both the data and design justice movements (Costanza-Chock, 2018; Heeks & Shekhar, 2019) as well as HCI more broadly (J. Carroll & Rosson, 2013; Chamberlain et al., 2012; Rogers, 2011; Rogers & Marshall, 2017; N. Taylor et al., 2013). The concept of sustainability asks that we consider what it takes for technologies to continue to be used and thrive once the designer's involvement in the project is over. Sustaining the data infrastructure not only involves the technical tasks of

staying on top of changes to the underlying platforms and API's, updating underlying datasets to include the latest data available and making usability improvements to the interface, but also includes social and institutional aspects of promoting it at various events, educating individuals and communities on how to use it and continuously negotiating the everyday politics of its constituent actors and networks. These were again tasks that fell under the purview of the WCA, but now had no takers. Designing of data infrastructures is therefore never really complete and should therefore continue through use (P. N. Edwards et al., 2007; Karasti & Blomberg, 2018). A justice oriented approach would require that we work towards sustainable, community-led and controlled solutions (Costanza-Chock, 2018) while preparing to respond to the variety of sociotechnical events that cannot be completely anticipated. Empowering data infrastructures are therefore those that lead the community towards action by revealing previously invisible perspectives and being continuously available.

5.5.3 Foster accountability and responsibility in and through the data

The question of accountability came up multiple times during the workshops. Participants wanted to know who at Georgia Tech they could hold accountable for the data that was being put out through the dashboard. They wanted to know what I had learnt from all the time I had spent with them and how I would be accountable to the community leaders and the dashboard going forward. Some even saw the dashboard as an accountability tool and wanted to use it to hold others individuals and organizations in the community accountable for the work they were doing and the impact they were promising to the residents. The fact that the interpretations suggested by data are always partial makes designers/developers more accountable for what they make (L. Suchman, 2002). Scholars

argue that when designing for accountability, the goal is to offer marginalized groups the data and tools they need to counter the dominance they face from other more powerful collectives (Dombrowski et al., 2016; Irani & Silberman, 2013). Costanza-Chock (2018, p. 11) argues that such accountability can be achieved by "the full inclusion of people with direct lived experience of the conditions the design team is trying to change".

The WCA claimed that their decisions for what data needs to be included on the dashboards was a result of their engagement with the matters of concern and with community leaders who were also experiencing them. But many of the decisions they took, as I described in Chapter 4, were also a result of the organizational, institutional and technical constraints that they found themselves in. Dean Royster, Sherri, Mackenzie and Katie were the public face of the WCA and the dashboard. But with the WCA gone, community leaders wanted to know who at Georgia Tech was the face of the dashboard, who they could report errors to and who they could contact to have their questions about the dashboard data resolved. There were also discussions about the distribution of this responsibility between the community and Georgia Tech. Were the community leaders collectively responsible to ensure that the data on the dashboard was accurate, or did this responsibility lie with Georgia Tech? The question of responsibility is related to that of accountability, in that while responsibility is task oriented, accountability is results driven. How would a data infrastructure ensure such responsibility and accountability in and through its data?

5.5.4 Disclose the context within which data are created

In order to create richer and more informative data stories, one needs to be aware of the context within which data are not only produced but also consumed (D'Ignazio & Klein,

2020; Vertesi & Dourish, 2011). Additionally doing justice to data stories asks that we always be conscious of not severing the connection between the story and the time/place from which the data are extracted (Parvin, 2018). Parvin offers the concept of *reciprocity* as one that draws attention to the give and take relationship that is enacted between the different actors involved in relaying stories through media. It points to the unequal gains, risks and responsibilities between the subjects of the stories and the storytellers. Parvin therefore asks that we be responsive to the vulnerability and uncertainty of the situation within which digital stories unfold.

Ambiguous and inaccurate data made many of the community leaders upset and question the provenance of the data. Loukissas (2016) reminds us that inaccuracies in data are a result of situated practices and assumptions in the production of data and are important signifiers of local conditions that should not be erased. "While one may never be fully aware of one's assumptions, disclosure describes the aspiration to be conscious of their potential effects and invite the viewer into exchanges with the designer, reflections about the visualization, and engagement with an issue" (Dörk et al., 2013, p. 4). Therefore, a more productive strategy to deal with data provenance might be to disclose the decisions that were taken in creating it through every step of the process. Attributing the work appropriately and being transparent about the labor that went into creating the data and its visuals would help to build credibility and trust with our audience. Such disclosure allows one to stay with the problem, which might lead to further insights about local data practices (D'Ignazio & Klein, 2016; Dörk et al., 2013). Additionally, Bardzell (2010a) asks that we also disclose the process through which software makes suggestions for its users, so users are able to discern the assumptions that

the software is making about them. When applied to data, such a strategy asks that the infrastructure make visible any data cleaning, formatting or calculations that were done when creating the specific visual or data story. This helps users be critical of the data and not take it as given as it raises the veil of objectivity that tends to be associated with numerical data.

The workshops helped reveal how participants were proficient at considering the historical, social and political backdrops of the situations that shaped the data stories, but unequally skilled at deciphering data and creating stories around them. Each of the participants brought their own situated and partial perspectives, which helped contribute to a collective data literacy (Frank et al., 2016). What they needed was a way to combine each of their contextual expertise and personal experiences with the technical expertise of finding data, analyzing them and creating data stories with the given tools. How can a data platform/infrastructure not detach data from their context, acknowledge their situatedness and not dilute them when developing alternate perspectives and interpretations? Additionally, how can a data platform/infrastructure disclose and maintain data provenance, while supporting community leaders in their continued use and reinterpretation of the context within which data are extracted and used? These, as I explain in the next chapter, are questions I considered when reimagining the Westside data infrastructure.

These localized values of data justice, along with other needs I identified during the workshops guided my reimagination of the Westside community's data infrastructure. I describe this process and the theoretical foundations which helped me reimagine the data

infrastructure in the next chapter. For now, I return to the workshop guidelines that form the second contribution of this chapter.

5.6 Guidelines for Infrastructuring Data Literacy within Communities

Participants in the first round, who attended four workshops over four weeks seemed a lot more willing to participate in the activities than participants in the second round, who attended four workshops over a two-week period. The breakdowns that the participants experienced in the second round of workshops definitely contributed to their frustration, but I wonder if the frequency and number of workshops they were asked to attend were also a contributing factor. Participants in the second round of workshops were also a lot more vocal about the shortcomings of the dashboard and the workshop activities, which is the reason many of the quotes presented above come from participants of the second round of workshops. Another difference I observed but did not collect empirical data about was about the average age and digital literacy skills of the participants; participants in the first round of workshops were on an average older and required more help with technology than participants in the second round of workshops. The nature of structuring such workshops in the wild makes it challenging to account for all of these variables and control them. Going forward, design researchers interested in organizing such workshops would be well served by knowing that there are a lot of factors that one might not be able to control for during such workshops, which is also the nature of design based research (Barab & Squire, 2014).

The success of any workshop depends on how well the facilitator is able to marshal their resources and think on their feet during the activities and discussions. While I did my best to follow the workshop facilitation guide, there were instances where I had to

deviate from the script in order to accommodate the ongoing discussions. Moreover, as the preceding conversations indicate, there were instances where the community leaders felt like I was being manipulative by intentionally presenting irrelevant and inaccurate data. While I did my best to convince them that I intention was not to deceive them but to unpack the infrastructural black boxes underlying the dashboard, I am not sure persuaded they were. This is where doing the work it takes to gain access, build rapport and trust, as described by LeDantec & Fox (2015) becomes crucial to working with communities. While I did volunteer with several organizations, attend public events and meetings and spend time in getting to know the neighborhood and its issues, I wonder if working with the participants more directly rather than going through the WCA would have allowed me to build more of a personal relationship with the community leaders. Such a relationship would foreground my research intentions rather than those of the WCA, which were often confused with those of its parent university.

The data stories these community leaders created always started from their personal experience with the related matters of concern. These findings compliment the results S. L. Erete et al. (2016) have reported in their study of data storytelling practices within non-profit organizations (NPO's), where they found that NPO's combine statistical facts with personal stories to create more impactful narratives. The way these civically engaged leaders communicated with and understood data was therefore different from the process most data analysts follow to let data speak for themselves, which includes discovery, wrangling, profiling, modeling and reporting (Kandel et al., 2012). These leaders relied first and foremost on their expertise that resulted from their years of

experience dealing with community concerns, and then connected it to data to strengthen their stories.

This also points to the need to consider data literacy as an aspect of the community, rather than that of an individual. (Frank et al., 2016) argue for "data literacy as a property of a community as opposed to an individual, with members of the community making different contributions. So that the presence of some people who can find data, some people who can manipulate it, and some who can present the result might constitute data literacy for that community". This would imply that we cannot expect everyone to be equally adept at all aspects of data access, interpretation and use. It is more reasonable to consider data literacy as an aspect of the collective where individuals take on specific roles in specific situations.

When the activities had data that was not related to the community, some participants seemed disengaged. I compensated for this by restructuring activities to include data that was directly relevant to the community leaders' concerns. But when the issues were too familiar, participants seemed more interested in talking about the infrastructural issues related to the data rather than focusing on the activity or creating data stories. While these discussions took time away from the activities, it was also when participants were at their critical best and brought up issues that would otherwise be overlooked. Design researchers might therefore consider offering a similar combination of activities that involve familiar and unfamiliar data.

I chose to structure the data infrastructure literacy workshops over the dashboard as the dashboard is part of a larger research agenda that studies data from an infrastructural perspective. Although the dashboard offered a window into the community's data

infrastructure, it is not necessary for such workshops to be based on data dashboards only. This is because even simple data collections result from interactions between a network of infrastructural black boxes that need to be unpacked if one is to truly understand how data are made. Any collection of community data or a data visualization can serve as the basis on which to build a community's data infrastructure literacy.

Based on the discussion so far, I present the following as an initial set of recommendations that might guide design researchers when organizing data literacy workshops with minoritized communities.

1. Invite participants for whom the data and the issues it represents are *matters of concern.*

All data are approximations and do not offer a complete picture of the world they represent. Identifying their shortcomings requires that they be analyzed by experts. The first guideline therefore asks to ensure that the workshop participants are experts in the issues that the data represent. One does not need expert data analysts. Instead one needs individuals who are familiar with the local contexts from which data are extracted. They need to care enough about the data or the underlying issue that they consider it a matter of concern (Latour, 2008; C. Le Dantec & DiSalvo, 2013).

2. Take a *critical pedagogical approach* that aims to contextualize the workshop activities with specific matters of concern that the participants identify with.

This means selecting data and activities that match the context of learning with the context of use. The critical pedagogical approach as advocated by Paulo Freire, calls for establishing a dialog between the teacher and the pupil in order to contextualize the learning to the pupil's needs. The exploratory workshops allowed me to identify key issues and select datasets, visuals and activities that were relevant to the target audience. I used the data storytelling framework to structure our workshop activities since it is predominantly how non-profit's tend to use data, either when advocating for change or raising funds (C. Clarke, 2009; Erete et al., 2016). I did this, so participants could draw parallels between the knowledge and skills they learned at the workshops and the work they did every day.

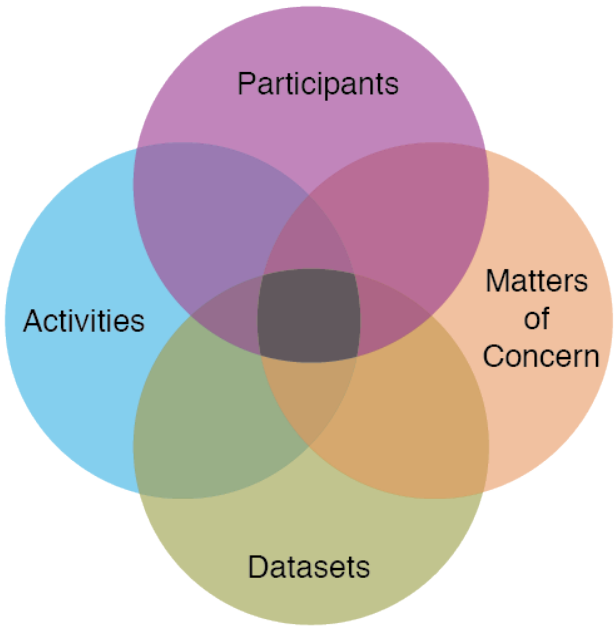


Figure 24: Visualizing activities at the intersection of the available datasets and matters of concern relevant to the participants.

One way to visualize the relationship between the participants, datasets, activities and matters of concern is shown in the Venn diagram above. This overlapping area represents

the subset of activities and datasets representing specific matters of concern that our participants care about and are apt to be included in a data infrastructure literacy workshop. The other overlapping areas meet two or three of the criteria and might fail to generate enough interest with the participants. The specific matters of concern, participants and datasets will need to be identified in an organic manner by establishing a dialog with the target community. The NNIP (National Neighborhood Indicators Partnership, n.d.), databasic.io or the Data Culture websites have a wealth of resources for possible activities and data one could use in such workshops. The datasets for any of these activities can be easily replaced with ones that the participants are more familiar with.

3. *Be open to an infrastructural imagination,* where participants reimagine the data infrastructure by critically analyzing its black boxes to reveal embedded barriers, opportunities and values.

Being *open to an infrastructural imagination* (Bowker, 1994), means asking that organizers encourage participants to reflect on the different socio-technical obstacles that need to be overcome in order to use the data. One way I did this was to introduce data and computational friction into the activities, which revealed the specific individuals, organizations, protocols, practices and policies that are implicated in the production of data. This also brought attention to values like trust, disclosure and responsibility that were a priority for the participants. The purpose of an infrastructural imagination is to foreground the often ignored background elements within the infrastructure so they can reassembled from scratch if needed (J. Gray et al., 2018).

4. *Focus on moments of breakdown* and be open to discussing inaccuracies and errors in the data infrastructure.

Given that the workshop participants engaged in their most critical discussions when they encountered inaccuracies and classification errors, I argue that the overall goal of data infrastructure literacy workshops should be to offer a platform where these moments of breakdown can come to the fore. These, as I have stated before, are important signifiers of situated data practices and should not be erased or ignored.

5.7 The Human Infrastructure of Civic Data Infrastructures

I conclude this chapter by returning to the human infrastructure that I interacted with during this phase of my research. The workshops offered a space for me to reconsider these values and continue to design after design (Ehn, 2008) by surfacing the different barriers to use, inaccuracies and mal-alignments (S. Star, 1999) within the data infrastructure. The workshops also helped me identify the infrastructural features that are crucial to ensure that the dashboard can meet the community leaders' data needs. These community leaders form a crucial element of the human infrastructure, as they are the local experts for other actors (community champions, transient workers, institutions, and organizations as identified in Chapter 4) who wish to engage with the community. These leaders are embedded in the local context and often work with these other actors to advocate for their perspective on what it is like to live in these neighbourhoods, which is often different from what external actors like media organizations and researchers portray. The perspective of such external actors was, in several instances during the workshops, at odds with that of the community leaders. These external actors therefore significantly influence how these community leaders like to use data and the narratives

they choose to tell with them. The external actors therefore represent an additional layer of the human infrastructure of the Westside's civic data infrastructure. I add *community leaders* and *external actors* to the human infrastructure table I introduced in the previous chapter. I will continue to add to this table as I describe the next steps of my research in subsequent chapters.

Table 5: Elements of the human infrastructure

Infrastructural Elements	Humans that Put the Infrastructural Elements Together	Type of Human Infrastructure
Advocacy efforts which include organizing meetings, programs, events, festivals, workshops etc	Community Leaders	Community Leaders
Media organizations, other researchers, communities.	Individuals representing these collectives	External Actors

In this chapter, I described the approach I took to organizing data literacy workshops with the Westside community leaders. Taking a critical pedagogical approach, my goal was to use the dashboard and associated activities to expand the notion of data literacy to data infrastructure literacy. Such a data infrastructure literacy led to an infrastructural inversion of the dashboard and raised questions about "social, historical, cultural and political settings" in which data are created and used, in order to highlight the different ways such infrastructures influence data (J. Gray et al., 2018). The breakdowns that led to this infrastructural inversion also revealed different contexts of use, unmet needs, desired values, and barriers faced when accessing and using data on the dashboard.

I compiled these observations into a set of local values of data justice and workshop guidelines that design researchers can use when organizing data literacy workshops within other communities. Such a method of using data literacy workshops to elicit local values is a response to Le Dantec et al's., (2009) call for more prescription in the local conceptualization of values that can guide the design of technologies. It also offers a response to L. Taylor's, (2017) call for a more local conceptualization of data justice values in order to operationalize her data justice framework. Each community is different and knowing what values they prioritize will allow us to build data infrastructures that have a better chance of embodying them in their design. I build on the findings of this infrastructural inversion in the next chapter to elaborate on issues of in/justice that I observed during the workshops.

CHAPTER 6. INFRASTRUCTURING FOR DATA EQUITY

In not following participatory and democratic design principles, the WCA dashboard furthered many of the injustices that the Westside community were already grappling with. In this chapter, I describe an experiment in which I used the insights from my research to counter the injustices inherent in the WCA dashboard and in my facilitation of the data literacy workshops. The process of infrastructuring towards socially just outcomes serves as a case study in which I democratize the process of building, using and maintaining a data infrastructure within resource constrained environments, while adhering to the data needs and values identified by its community leaders. The case study therefore offers one possible response to my research question, *What are the local values and infrastructural arrangements that are required to build, use and maintain equitable data infrastructures that enable communities to benefit from the publicizing of data through dashboards?*

Several frameworks have been suggested in order to translate values into design (Flanagan et al., 2008; Friedman et al., 2009). But as Parvin et al, publishing as JafariNaimi et al., (2015) argue, the work of identifying values does not precede design, but helps in further developing the design situation and the required response. While I separate the identification and specification of these values from their subsequent operationalization into separate chapters, this process was in reality one extended design intervention that involved wrestling with these values and negotiating the best way they could serve the Westside's civic data infrastructure. It was through this redesign that I attempted to counter many of the injustices that the community leaders had experienced

through the dashboard and its data. My hypothesis in this redesign was that in order to achieve data equity, the Westside community's data infrastructure needed to (i) support data infrastructure literacy, (ii) empower the community towards action, (iii) foster accountability and responsibility in and through the data, and (iv) disclose the context within which data are created. These values, along with the needs identified in the previous chapter guided my redesign of the Westside community's data infrastructure towards more socially just outcomes, which in this case was more equitable participation in accessing and using their data. I approached this redesign through the lens of infrastructuring, which Grönvall et al. (2016) have argued allows one to engage in community based PD while considering a dynamic set of values.

Infrastructuring as a concept has taken on greater significance in the last few decades (Korn et al., 2019). I review some of this literature in the field of PD here and also describe the participatory approach I took in infrastructuring the Westside community's data infrastructure. This process of participatory infrastructuring led me to identify data intermediaries as another element of the human infrastructure that needs to be considered when infrastructuring data infrastructures for communities. I describe the manner in which I engaged with these data intermediaries and conclude this chapter by discussing the results that came out of this process.

6.1 Infrastructuring and Long-Term Design Issues in PD

Infrastructuring advocates for an asynchronous design process, one that is separated in time and space from what the designers engage in at design time (Ehn, 2008). Rather than finalizing all the design decisions at design time and only evaluating its use later on, it asks that we focus on creating environments in which design can continue after design;

where we see every usage scenario as an opportunity for design. Le Dantec & DiSalvo (2013) argue that PD offers the appropriate set of methods for infrastructuring towards the future Infrastructuring for them, involves moving away from 'designing for use' to 'creating fertile ground to sustain a community of participants' or *publics* around specific matters of concern.

This idea of a public was first proposed by John Dewey in his book *The Public and its Problems* (Dewey, 1927) and has undergone several revisions since then (Calhoun, 1992; Fraser, 1990; Habermas, 1991; Lippmann, 1993). The idea of publics is relevant here because it refers to a unique set of individuals who come together under specific circumstances as a response to proximate or future matters of concern. It is through the process of infrastructuring that these publics discover their *attachments* to resources that allow them to develop relevant sociotechnical responses to these concerns. Le Dantec & DiSalvo (2013, p. 255) further point out that "an important aspect of infrastructuring is recognizing that those attachments are dynamic; they will change, often in unanticipated ways". Three ideas that they offer are relevant to my implementation of infrastructuring here. First, they suggest that the process of infrastructuring need not exclusively be concerned with the design of future systems and the discovery of unknown issues, but can also inform the design of more proximate useful systems and their issues. Second, infrastructuring for them is a participatory process that helps initiate a public (through a technological intervention in their case) and sustain it over time by developing attachments to specific socio-material resources. Third, they point out that infrastructuring can be used to respond to entrenched authority or other networks of control, expertise or influence, which is what makes it relevant to addressing issues of

social justice. The potential for such a sociotechnical infrastructuring to the formation of relevant publics and long-term intervention in issues of social justice is what makes it relevant to infrastructuring the Westside's data infrastructure as well. I offer a couple of examples of how infrastructuring has been used to address social justice concerns next.

In infrastructuring a food democracy, Prost et al. (2019) document their process of setting up a local food hub in a socio-economically deprived neighbourhood of England. Similar to the arc that my dissertation took, setting up the food hub started with a focus on setting up the infrastructure and moved onto infrastructuring of environments in which the concerned publics take on the responsibility to grow and sustain the infrastructure in the absence of the design team. In order to describe their process of infrastructuring, they draw on the notion of Things as socio-material assemblages of humans and non-humans around a shared matter of concern (Bjögvinsson et al., 2012; Latour, 2005). Things like design coalitions and the reconfiguration of existing technical platforms like social media, email, newsletters and telephones supported the formation of publics that drove the community towards social innovation in the civic food network. They use the term *guerrilla infrastructuring* to refer to this process of reconfiguring existing platforms, people and practices to meet immediate design needs.

The term *guerrilla infrastructuring* was first put forth by Vlachokyriakos et al. (2018a), which they defined as the "the infrastructuring of alternative economies and more generally spaces of designed social innovation through creatively using resources from, and resourcefully negotiating synergies with, the mainstream economy and centers of power." They refer to it as a form of "irregular, opportunistic, adaptable, responsive and decentralized strategy of infrastructuring social innovations and reform that stand in

contrast to existing power and economic configurations". They use the concepts of strategies and tactics, first developed by De Certeau & Mayol (1998) and then exemplified by DiSalvo (2009) in the construction of publics to describe their process of building coalitions and assuming different design roles in service of the Solidarity Movement in Greece. Furthering the Solidarity Economy in Greece required that they move between the polarities of expert and diffuse design (Manzini, 2015) in order to reconfigure existing systems, localise existing platforms, develop bespoke systems and scale them out in an effort to infrastructure for social innovation.

Agid (2018) extends PD's concern with the political to suggest that the aims with which we choose to infrastructure with are critically important to what it means to do infrastructuring with different communities. Agid situates her work in the 'local' infrastructures of Critical Resistance, a US-based group that is working to abolish the practices and effects produced by the US Prison Industrial Complex (PIC). In doing so, she acknowledges the relationality of infrastructures (S. Star & Ruhleder, 1996), her own situatedness within them (Haraway, 1988), and the interpersonal process of conflict and negotiation that is involved in navigating and locating accountability within this complex infrastructure (L. Suchman, 2002). Agid's position as a designer shifted dynamically with the ideas, experiences, locations, artifacts, relationships, needs and goals for which they were infrastructuring. In designing for local infrastructuring in this way, Agid's material design interventions are located in the present, with an eye towards a future in which her practices of infrastructuring can counter the harm produced by the PIC. In doing so, she argues for using PD to explore the "intersections of local, large-scale and imagined infrastructures, while emphasizing the importance of building relational practices to do

infrastructuring in meaningful ways grounded in ongoing day-to-day work with people" (Agid, 2018).

This concern for the local/situated practices and dynamic relationships within infrastructures also comes forth in what Karasti & Syrjänen (2004) refer to as *artful infrastructuring,* which is similar to the concept of guerrilla infrastructuring described previously. Building on L. Suchman (2002) idea of 'artful integration' of technologies and infrastructures by S. Star & Ruhleder (1996), Karasti & Syrjänen (2004, p. 28) draw attention to how the "blurring of boundaries between use, tailoring, maintenance, reuse and design, as well as attention directed to local, situated everyday practices with technologies, forms design as artful infrastructure processes which are tentative, open and flexible".

My approach to infrastructuring builds on these concepts of infrastructuring, and involves end users in the design of technologies that they use. This form of participatory infrastructuring treats the object of design as boundary objects through which new object of design can emerge through use. Such an "ongoing and continuing processes of creating and enacting information infrastructures" (Karasti & Blomberg, 2018) that is entrenched in the political aspects of participation in an attempt to sustain a public through the infrastructuring of relevant socio-material resources is what I adopted during this phase of reimagining the Westside community's data infrastructure.

Additional guidance came from the more recent critiques of PD, in which scholars have criticised the field for being obsessed with small issues which involve democratizing participation to serve the usability of products rather than engaging with larger societal issues that engage with the political aspects of design (L. Bannon et al.,

2018; Bardzell, 2018). Bødker & Kyng (2018) argue that we need to move beyond the here and now and take a longer-term perspective when thinking about end user participation in designer led research projects while also paying attention to the benefits accrued to each of these stakeholder groups. It is important to consider who is participating, their roles, agencies, capabilities, conflicts, politics and ethical issues involved in bringing about their preferred visions. Rather than relying on the expert designer, the alternate visions must be produced by stakeholders to whom these visions matter the most. Such a practice also helps build a culture of local accountability and sustainability within the stakeholders of the project.

In reimagining PD, scholars have suggested we reinterpret early PD projects like NJMF (1971-73) (Nygaard & Terje Bergo, 1975), Demos (1975-79) (Ehn & Sandberg., 1983; Sandberg, 1979), DUE (1977-80) (Kyng & Mathiassen, 1980) and Utopia (1981-84) (S Bødker et al., 1987) as models for how PD ought to be done and formulate a version of PD that is borne out of the neoliberal context in which we live today where trade unions no longer hold as much power as they used to (Bardzell, 2018; Susanne Bødker & Kyng, 2018). How would a reimagined PD be able to scale to accommodate the IT infrastructure larger organizations, cities and even countries? What new methods and roles would designers need as PD takes on global and systemic issues like climate change, immigration and the rise of authoritarian governments? Such a reimagined PD would need to focus on not just the small, but also the larger and longer time horizon in which such issues unfold.

These concerns with sustainability and scaling of PD are also shared by Smith & Iversen (2018), who suggest three dimensions to create sustainable social change via PD.

These dimensions stem from Smith and Iverson's concern that contemporary PD projects tend to favour pragmatic design solutions rather than engage deeply with the political aspects of empowerment, democracy and voice, which were PD's grounding values (Kensing & Greenbaum, 2013). The three dimensions, *scoping, developing and scaling* are not necessarily mutually exclusive and offer a renewed focus on PD's core values. *Scoping* refers to 'Stage 0' of the project, which starts before traditionally reported PD projects and involves exploring the field with diverse stakeholders in order to surface the 'protagonist communities' who have relevant participants and research questions. In *developing*, the focus is on building understandings and frameworks of new digital practices rather than producing technological artefacts. These efforts are then *scaled* by scaffolding appropriate networks, frameworks and visions for stakeholders who can take the project towards sustainable social change. In doing so, they hope to move beyond mere 'user participation' and include longer term participatory infrastructuring Bødker et al. (2017), where the focus is on development of sociotechnical systems and networks of actors that can have societal impact.

Bardzell (2018) has suggested a feminist utopian reimagining of PD theory, methodology and practice that allows practitioners to pursue the democratic, political and social justice agendas that PD was originally intended for. In proposing a feminist utopian PD practice, Bardzell asks that we combine PD's methodological strengths that democratize design within politically conflicted situations with feminist utopian practices of accommodating a plurality of voices, resisting relativism and being open to imagining better futures rather than attempting to define them.

Bødker & Kyng (2018) suggest eleven elements of a reimagined PD. The first five are focussed on the political dimensions of a new PD practice and the last six are meant to support a more action-oriented approach to PD. I will briefly outline them here and then describe how my process of reimagining the Westside community's data infrastructure aligned with these suggestions.

The five suggestions about enhancing the political dimensions of PD are

1. PD should address areas where dramatic, potentially negative changes are under way.
2. Partners are major drivers of engagement and action.
3. Researchers play a dual role of researchers and activists.
4. Researchers and partners cooperate on a vision for high and lasting impact to counteract potentially negative developments.
5. Impact must be safeguarded by developing democratic control mechanisms.

The six suggestions which focus on supporting action through more technological content are

6. High technological ambitions are necessary in order to influence our technological future.
7. Deployment of working prototypes is an important tool for creating impact and experiences, and for sustaining a PD initiative over prolonged periods of time.
8. Alliances with other actors beside the direct partners are an important part of creating and sustaining impact.
9. We need to consider how to develop and use ideas and results on a larger scale.
10. We need to understand the issue of success vs failure.

11. New PD is also action research.

These guidelines for a reimagined PD that focuses on the political and longer-term issues also served as guidelines in my participatory infrastructuring of the Westside community's data infrastructure. I describe the specifics of the participatory infrastructuring process next.

6.2 Participatory Infrastructuring of the Westside Data Infrastructure

The first step in the participatory infrastructuring of the Westside data infrastructure was to focus on the scenarios through which the community leaders would engage with the data infrastructure and translate them into requirements. The table below translates all of the issues outlined in the previous chapter into specific features to be supported in the reimagined data infrastructure. The first column lists the aspects about data and the dashboard that community leaders (CL's) found issues with. The second column translates these requirements into features on the dashboard/ data infrastructure (DI) that needed to be rethought.

Table 6: Translating observed infrastructure issues to requirements.

	Outstanding issues with the data dashboard/infrastructure	**Features to be supported in the reimagined data infrastructure**
R1	CL's do not have the capacity to analyze data themselves. They need help from others in doing so.	Capacity to collaborate with experts who can analyze and present data in relevant formats.
R2	CL's are extremely busy and do not want to attend additional workshops to	Tools offered should be intuitive to use and/or offer sufficient online

	learn how to use data.	documentation for support.
R3	CL's need data about hyperlocal issues like where to start community gardens, use of recreational facilities, youth programs etc.	Capacity collect and include relevant hyperlocal data.
R4	CL's need to know who is responsible for the accuracy of the data and where errors can be reported.	Transparency about who is in charge of the data and how they can be contacted to report any issues.
R5	CL's need to be able to talk back to data. Their personal perspectives about data need to be a part of the data stories.	Capacity to annotate/modify existing data stories or create new ones if required.
R6	CL's need help monetizing data or using data to apply for grants or funding.	Share information about relevant funding and guidance on how to apply for them.
R7	CL's want to use the dashboard to hold those who are represented accountable for what they have done.	Data sources and processes through which they were obtained need to be clearly specified.
R8	CL's should be able to verify the accuracy of the data themselves.	Build collective capacity so the CL's can verify data themselves.
R9	Data infrastructure needs to be available and usable with the latest data at all times.	The data infrastructure needs to be preserved and maintained beyond the project's funding cycles. Availability

		of the infrastructure should not be subject to individual institutional politics.

Building an entirely new tool or platform from scratch was out of the question, as this would introduce some of the same issues around maintenance and sustainability that were concerns with the current dashboard. I looked for existing software tools and platforms that would support the identified data needs and concerns of the community leaders. I found that using a combination of Google's Data Studio, Google Drive and Google Groups met most of the identified needs.

A key consideration here was making sure that the platforms I was putting together were replaceable based on how they were used by the participants. I chose the platforms I did because they worked well together, allowed for collaboration to be controlled and authenticated, were well supported and did not require too much technical knowledge to setup and run. Being made and supported by a company like Google meant that these platforms would rarely go offline and would always be available to its users, thereby satisfying R9 (*Data infrastructure needs to be available at all times*). Additionally, Google also makes sufficient documentation available for its products and has a huge user base. This feature satisfies R2 (*CL's are extremely busy and do not want to attend additional workshops to learn how to use data*) which should make finding product support and learning to use the different tools a lot easier.

The Data Studio platform allows multiple users to work together in exploring data and building reports and makes this entire process quite transparent. This is an advantage for

those looking to share the process through which the data were manipulated when building the report or just ask for help from others, thereby supporting collaboration and satisfying R1 (*CL's do not have the capacity to analyze data themselves. They need help from others in doing so*). Data Studio does not require any software or libraries to be installed on the local machine as everything runs within the browser. All one needs to use it is a Gmail account, which many of the community leaders seemed to have (this is based on the email addresses participants used to sign up for the workshops). It is managed by similar account creation, authentication and access features that are common to other Google products. Similar to products like Google Docs, Sheets etc., Data Studio specifies who the owner of the file is and allows them to create shareable links and assign read/write access to their data reports. This feature satisfies R4 (*CL's need to know who is responsible for the accuracy of the data and where errors can be reported*) as one always knows who the owner of the report is so they can be contacted in case of any clarifications or errors.

Within Data Studio, datasets are associated with specific users and the reports they are used in. While these datasets can be shared with others, any changes one makes to the data types would also be shared with other users. All of these data manipulations are visible to collaborators with read/write access to the report. Such users can potentially make a copy of the report and modify it to suit their specific needs, thereby satisfying R5 (*CL's need to be able to talk back to data. Their personal perspectives about data need to be a part of the data stories*). Each of these reports can be based on single or multiple datasets that can be joined based on specific rows or columns of data available. Data Studio also has a wide range of data connectors that allow users to import data from a

wide variety of applications. These include 18 that are built and supported by Data Studio, 163 partner connectors and 3 open source connectors. This makes it easy for users to connect to data in a variety of formats or databases from basic csv, Excel or MySQL to other proprietary data on platforms like Amazon, Reddit, Comcast etc. This feature satisfies R3 (*CL's need data about hyperlocal issues like where to start community gardens, use of recreational facilities, youth programs etc*) so community leaders can find and upload their own datasets which might be made available to them in a wide variety of formats.

I chose to use Google Drive as the platform to store and distribute community datasets. I created a hierarchical folder structure for all the relevant data I could find and included notes about where I downloaded them from or who I contacted in order to get them. I did this for data I downloaded from Neighborhood Nexus and the Atlanta Regional Commission, as well as the data I received from individuals at other organizations. This helped satisfy R7 (*CL's want to use the dashboard to hold those who are represented accountable for what they have done*) which required knowing where the data came from. I created links to these folders and shared them with the rest of the cohort via the Google Groups platform.

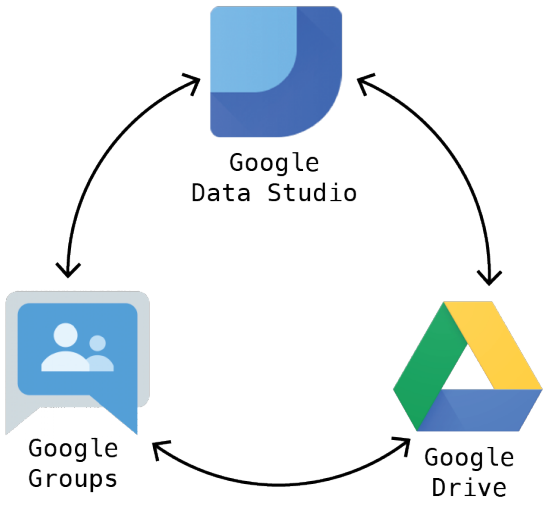


Figure 25: The reimagined civic data infrastructure has been prototyped with Google's Data Studio, Groups and Drive platforms.

Google Groups makes it easy to manage email subscriptions to large groups of users. Once users sign up, they can send emails to everyone in the group without having to worry about managing and finding individual email addresses. Setting up this group was my attempt at satisfying R6 (*CL's need help monetizing data or using data to apply for grants or funding*), by helping its members find and share relevant information about funding, local events, new data sources or missing data a lot easier. This bricolage of commercial off-the-shelf platforms helped me address most of the issues listed in Table 4. In order to build the capacity to work with data R2 (*CL's are extremely busy and do not want to attend additional workshops to learn how to use data*) and R8 (*CL's should be able to verify the accuracy of the data themselves*), I complemented these platforms with the Communities Who Know (CWK) Data Ambassadors program.

6.2.1 The Communities Who Know Data Ambassadors Program

The goal of the CWK ambassadors program was to identify individuals who had the time and interest to serve as data ambassadors for the CWK and others in the community.

These ambassadors would work closely with CWK and serve as liaisons between the university and the Westside for data related matters. The goal was to build enough expertise across the group of data ambassadors, so they could maintain and sustain the entire data infrastructure themselves with minimal support from the university (*R9 - Data infrastructure needs to be available and usable with the latest data at all times*). At the start of the program, I would work individually with these data ambassadors and support them as they acclimated themselves to all the data platforms I had put together. They would use the platforms to find data and create stories that were relevant to their concerns. The program would run initially for three months, after which we would reconvene to take stock of where we were and what needed to be done next. My involvement and support would reduce over time, as the ambassadors would serve as points of contact for all data related questions to the rest of the Westside community.

I presented all of this information and the different platforms to the 14 attendees who were present at the launch of the program and concluded by soliciting ideas for the kinds of data stories participants would like to create. Four of the attendees Casey representing Historic Westside Gardens, Roy representing NPU T, Hailey representing the Historic Westside News and Julia representing Saving Our Sons & Sisters International (SOSSI) suggested the following as stories they were interested in finding more data about. I summarize these stories in the table below and offer more details about them in the subsequent section.

My intention in detailing the specifics of how I went about searching for the data, even in the instances where I failed to find any, is to draw attention to the invisible labor (Crain et al., 2016) that is involved in the process. The first three data stories I describe

below are only meant to illustrate the kinds of actors, networks and labor that are involved in finding community data and the barriers I faced in doing so. This invisible labor involved in working with community leaders to gather the data for all the stories further highlights the importance of developing working relationships with data intermediaries, which is a point I take up further in section 6.3.

Table 7: Data requested by participants at the data ambassadors program launch.

Data Story	Data Needed	Data Available
Organizations doing impactful work in the Westside	Funding received, programs organized, local impact in the community	Needs to be collected.
Impact of food insecurity on children's school performance	Data about food deserts, food stamp usage and school performance within each of the Westside neighborhoods.	Needs to be collected.
Respond to G-DOT's Renew Atlanta $400m shortfall	Number of residents in their neighborhoods, number of seniors, their gender, racial makeup, family status (married, single, parents etc.) and vehicles per household along the corridors that include RD Abernathy, Cascade Ave, Campbelton Rd, & JE Lowery	Data available on the CWK dashboard and American Community Survey, but it is not the latest.
How are job training programs they stabilizing families?	Workforce training data	Data available from state and county departments.

6.2.1.1 Organizations Doing Impactful Work on the Westside

Casey had been living in the Westside for about 10 years and offered video recording and editing services for many of the events on the Westside. She was also part of the non-profit called Historic Westside Gardens that was helping setup community gardens in different neighborhoods around the Westside. Casey was interested in using the dashboard as an accountability tool and wanted to know which organizations were doing impactful work on the Westside. She wanted this to be a publicly available resource she and others could use to hold these organizations accountable.

No such list was publicly available at the time. One starting point was the same Blank Foundation website from which the visualization I presented in Chapter 5.2.2. The foundation, which is the philanthropic arm of Arthur M Blank, the founder of The Home Depot, has invested millions of dollars in local foundations and maintains a list of organizations they have given money to on their website. I shared this list with Casey to get her opinion about what she thinks of the impact these organizations had on the community. Not being able to define precisely what we meant by impact, Casey admitted that she didn't know of all of the organization on that list. She suggested reaching out to her contacts and asking them to suggest names of organizations that they thought were doing impactful work on the Westside. Asking Casey to send the email would improve the chances to receiving a response, since I did not know as many people in the community. I was also wary of reaching out to people I did not know personally and could not compensate for the time they were devoting to my project.

I was copied on the email Casey sent to her contacts, and they responded with a list of 25 organizations. Some of these organizations overlapped with the list of 1765 entries

that were available on the foundation's website. Not knowing how Casey wanted to proceed with this list, I complemented its entries with additional information I could find about the organizations like the programs they run, the funding they received and the impact they had on the neighborhoods. I formatted this information in an excel sheet and shared it with Casey and everyone through Google Drive.

6.2.1.2 Finding data to counter G-DOT's 400mn shortfall

Roy, who was from NPU T wanted to gather data about transportation patterns in the Westside neighborhoods. He described to me how GDOT was experiencing a $400 million shortfall and was in the process of reprioritizing their development efforts based on neighborhoods that needed it the most. He was afraid that the money that was promised for the development of their community streets would be redistributed to wealthier neighborhoods in the north. He needed data to demonstrate to the Georgia Department of Transportation (GDOT) that their neighborhoods needed their transit infrastructure to be improved. To do this, Roy asked for data about the demographics of the surrounding neighborhoods, how many cars they owned, the amount of traffic and where it came from.

Data about demographics and vehicle ownership is available from the Census/ACS and is on the CWK dashboard, but this is not the latest data. Much of the data about traffic patterns on specific streets is generated from phones and GPS (Global Positioning System) and is privately held by companies like Inrix, City Labs and Airsage. GDOT does have some data on their website (http://www.dot.ga.gov/DS/Data#tab-4) but much of this is undecipherable and not related to what Roy needed. There was no resolution in

sight as we could not find the data we needed in the few weeks that Roy had to respond to GDOT.

6.2.1.3 Finding data about food insecurity and their impact on children

At the program launch, Julia who was representing SOSSI (Saving Our Sons & Sisters International) suggested that she would be interested in finding out more about the impact of food insecurity on the performance of children in schools. I searched online and found a researcher at the University of Georgia who had collected data and done some work related to food insecurity in Atlanta. The USDA (United States Department of Agriculture) also has data about food deserts on its website. While research about the impact of food deserts on students and school performance exists, no data had been collected within Atlanta's schools, and specifically those on the Westside. In the weeks following the program launch, I emailed Julia multiple times to talk about this dataset and how we might be able to use. Not hearing back from her, I uploaded all the data I found to the shared Google Drive.

6.2.1.4 Workforce Training Data

Hailey, who had attended the previous workshops and was involved with multiple non-profits on the Westside was concerned about the job training that residents were receiving and if it was helping their upward economic mobility. She wanted to know if the workforce training programs were indeed helping residents attain higher wages and improve their standard of living. As houses became more expensive in the Westside neighborhoods, were these job training programs helping residents find the jobs they needed to afford living in these homes or were they having to move out?

While I started off working with all four community leaders mentioned here, the workforce data offered the most promise. This was because for one, Hailey and another community leader, Mark were willing to collaborate with me in building out this data story. Mark had attended one of the first exploratory data literacy workshops I had done and had consistently expressed interest in being involved in the development of the data dashboard. Second, a workforce training data module was already present on the CWK dashboard and was in line to be updated. All the module had was a static list of workforce training organizations that served the Metro Atlanta neighborhoods, with no actual data about their performance or other metrics. There was always interest in building out this module, but other dashboard issues always seemed to take precedence over it. The CWK ambassadors program presented an opportunity to build out this module in collaboration with the community. Third, the State of Georgia had the data we needed, and it was a matter of finding the right person to give us access to it.

In my quest for community data, I came across the Fulton County Performance Dashboard, whose aim is to promote transparency in the operations of the county government. The dashboard puts out data in the form of reports across six priority areas which relate to people's safety, economic opportunity, self-sufficiency, health, cultural enrichment and the government's fiscal efficiency. I met with one of their data professionals, Shawn, in the hope of finding the data I needed. Shawn in turn put me in touch with his colleague who used to oversee the workforce training data for the State of Georgia, who in turn helped me setup a meeting with the current Workforce Division Manager who then introduced me to his data professionals, Tara and Gary.

Traversing this chain introduced me to the several actors that are involved in administering and maintaining data about workforce training in the State of Georgia. The Technical College System of Georgia's Office Workforce Development is the administrator of WorkSource Georgia, the state's federally funded employment and training system which is working to connect talent with opportunity. These federal funds are part of a grant program called the Workforce Innovation and Opportunity Act (WIOA). WIOA funds are allotted to individuals or businesses and administered specifically through services geared toward helping disadvantaged citizens obtain meaningful employment. At a local level, WorkSource Georgia provides WIOA services across the state through 19 local offices, of which the Atlanta Regional Commission (ARC) is one of the biggest. ARC, which is #7 in the figure below, is also the local office that serves the Westside community neighborhoods.

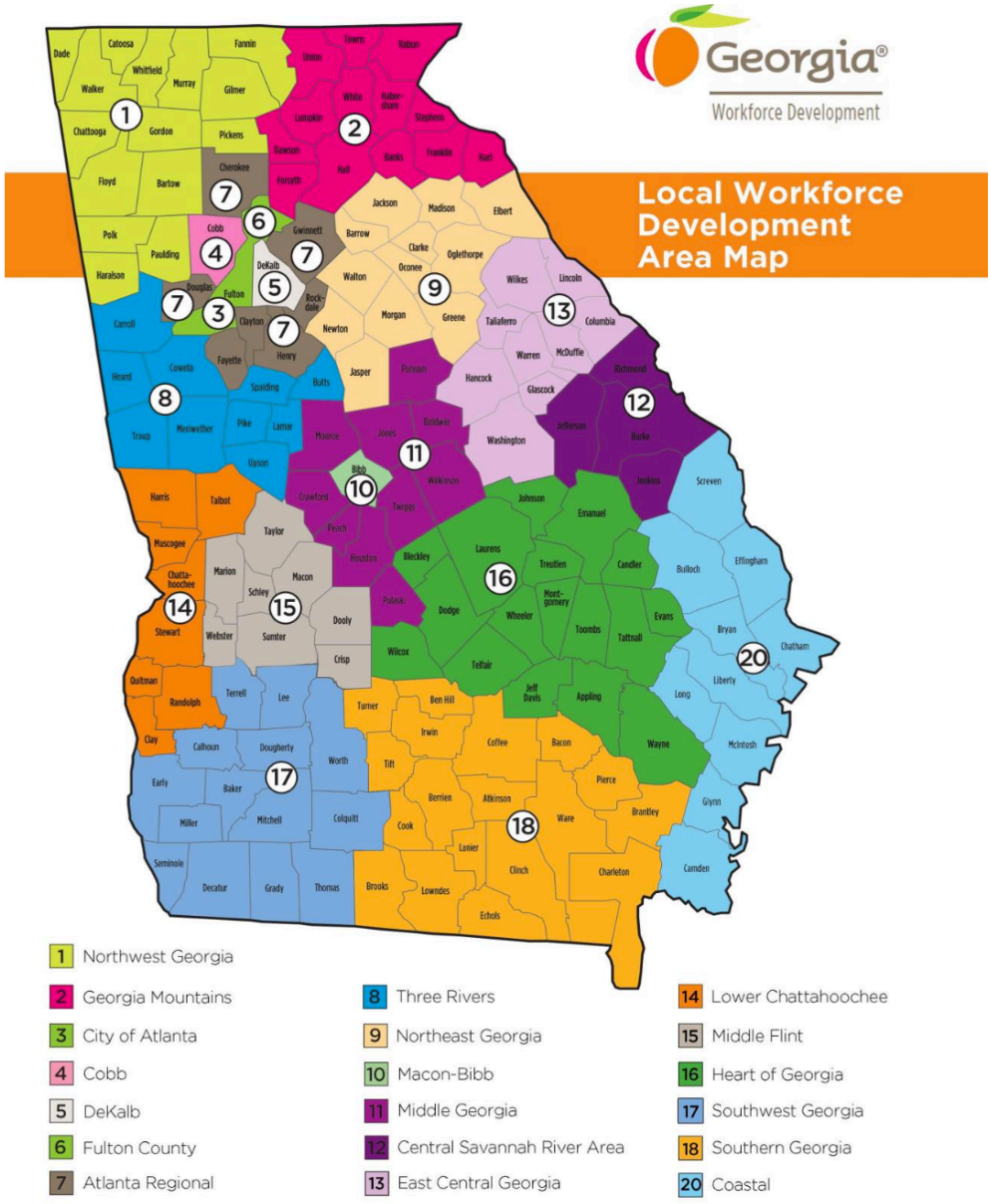


Figure 26: Local Workforce Development Area Map for the State of Georgia.

In finding and analyzing the workforce data, I was guided by the Civic Data Guide created by Loukissas, with assistance from D'Ignazio and Gradeck (Loukissas, n.d.). The guide offers a step-by-step process for students to think critically about how data are made in different settings and the process that is followed to do so. The first step in the

process is to pick a publicly available dataset and unpack its data *setting* (Loukissas, 2019, p. 2). This means investigating the source where the data came from and examining their standards, formats, typical uses, anomalies, audience, ethical issues and the like. For this first step, I chose to use a copy of the workforce training data I received from Tara and Gary, who are data professionals I met with at WorkSource Georgia.

In my first informal meeting with Tara and Gary, I explained what I was doing and described the data I needed. Based on my description, Tara sent me a dataset that included data about adult and youth participants in the WIOA program. I analyzed this data with Mark, who expressed interest in working with me after I announced on the Google Group that I had acquired the workforce training data and was looking for someone to collaborate with to do the analysis. I formally interviewed Tara a second time after I had analyzed the datasets and had a few questions about them. This interview lasted for about 70 minutes, which I recorded and transcribed for analysis. I interviewed a second data expert Aaron over email, as we were unable to find time to meet in person. Such a contextual interview of the people who are involved in creating, managing or using those data also forms the second step of the Civic Data Guide.

The third and fourth steps of the guide involve creating data life-cycle diagrams and comparative visualizations that allow you to explore and represent the data in a more visual manner. In the next section, I offer details on the step-by-step process I followed to analyze the data and create the visualizations with Mark, Tara and other data intermediaries in the community.

6.3 Thinking Critically about Data

6.3.1 Origins of Workforce Data

Tara is a part time contractor with the Atlanta Regional Commission (ARC). She provides technical assistance for training and case management systems (called the WIOA) to all the 19 local workforce training areas, of which ARC is one. Tara came to the ARC by way of the Georgia Dept. of Labor (GA DoL), where she had a similar job for almost four decades. She, along with many other permanent employees were laid off in 2013 when the Governor moved the WIOA training program from the GA DoL to the Governor's Office of Workforce Development. This move meant that the data from the GA DoL had to be converted to fit into the new Geographic Solutions (GS) database system that was being used. This was a messy conversion, as not many people had worked with the data and knew how it worked. Start was hired by ARC so she could help make sense of this data that had been converted into the new GS system, while also continuing to support ARC and the rest of the workforce training areas (in that order). It took her a good two years to create the conversion logic and fix all the errors in the data!

6.3.2 Working with Workforce Data

On a daily basis, Tara's job involves working with the workforce training data and communicating any issues she finds with them. She does this by writing SQL (Structured Query Language) queries into the case management system that Geographic Solutions (GS) have built. She asks GS for data maps or models that can visually show how the data are structured in the database but GS doesn't always have them. And even if they do, they're not always accurate. She's had to learn the data in the new system by studying the specs and doing little tests herself. GS continues to change the structure of the database "everyday", so there is always something new to learn. Tara also supports data requests

that come from individuals and other departments almost every day. She is not sure if anyone else in the organization handles these, so she basically does all of them.

In my first data request for Tara (before I interviewed her formally), I asked for data about the kinds of training programs that the workforce training centers were offering, who were the participants and what kinds of jobs they had before and after the training. Based on my description, Tara sent me a dataset that included data about adult and youth participants in the WIOA program since 7/1/2010 from six zip codes (30308, 30309, 30310, 30313, 30314 and 30318, which are Westside zip codes). The dataset contained 5,244 records in the adults dataset and 1,539 records in the youth dataset. Both datasets had the following 44 columns.

Table 8: Column names in the first iteration of the workforce training dataset.

STATE ID	ZIP CODE	VETERAN STATUS	RACE	GENDER
AGE AT WIOA PARTICIPATION	HISPANIC	SCHOOL STATUS AT APPLICATION	BASIC SKILLS DEF AT APP 0=NO 1 =YES	ENGLISH LANGUAGE LEARNER 0=NO 1=YES
EX-OFFENDER	YOUTH IN OR AGED OUT OF FOSTER CARE	DISABILITY	WORK HISTORY JOB TITLE	WORK HISTORY OCCUPATION
WORK HISTORY START DATE	WORK HISTORY END DATE	WORK HISTORY SALARY	WORK HISTORY SALARY UNIT	WIOA APPLICATION ID
PARTICIPATION DATE	PARTICIPATION FUNDING	OST PROGRAM	OST OCCUPATION	OST BEGIN DATE

OST PROJECTED END DATE	OST ACTUAL END DATE	OST COMPLETION STATUS	OJT OCCUPATION	OJT BEGIN DATE
OJT PROJECTED END DATE	OJT ACTUAL END DATE	OJT COMPLETION STATUS	CASE CLOSURE DATE	CREDENTIAL RECEIVED
DATE CREDENTIAL RECEIVED	JOB TITLE	OCCUPATION	HOURS PER WEEK	HOURLY WAGE
QTRTOTAL	START DATE	TRNG RELATED EMP	EXIT DATE	

6.3.2.1 Unpacking the data setting

When Tara sent me the dataset, she warned me that some participants might be listed more than once if they

1. Indicated more than once race
2. Received more than one credential
3. Were in more than one Occupational Skills (OS) or On the Job Training (OJT) training
4. Had more than one WIOA Application
5. Had more than one Work History record (some customers have quite a few)

Another anomaly that Tara pointed to in her email was that there were duplicates of several records, but each record corresponded to a specific service that the participant received. The GS system creates a separate record for each service that the participant avails, and that is how Tara also extracts it from the system with her query. For instance, one 'STATE-ID', 284063 was listed 11 times but had the same 'APPLICATION-ID'.

Each row corresponded to a different 'WORK HISTORY JOB TITLE' for that participant. Moreover, it seemed like many of these records did not have a value for OST (Occupational Skills Training). Did this mean that they did not participate in any training program? Why were their names even included in the database then?

As Mark analyzed this dataset, he noticed that one workforce training ID was associated with 56 rows and four different races, which made him distrust the data. He wrote me an email which read

> *The more I look at the workforce dataset (Adult-Dislocated Worker-Youth Participants Active from 7-1-2010 Forward in Specified Zip Codes - Adult Participants.csv) the less I trust it. I find it interesting that two columns "Work History Employer" and "Primary Employer" from your list above are missing from the file.*
>
> *Also, if you query for State ID = 324056 you will find that even the race data seems to be messed up. This person has 56 rows in the table and 4 different races.*
>
> *I hope the new data you get will be better. I had hoped to use this dataset but I cannot find any useful trustworthy info from it other than attendance/completion related filtering and now it appears that even for that the race numbers may be bogus.*

This email was in response to the dataset that I had shared with Mark, in which I had also specified the names of the columns that Tara said she would share with me. Mark was pointing out that two of the columns that were supposed to be in the file were missing. Tara might have overlooked those column names or put them in the email by

mistake. I made a note to ask her about that when I spoke with her next. The 56 rows might have been a consequence of the 12 different job titles and four different races that the participant reported, which Tara mentioned would lead to duplication of records. These discrepancies in the data led Mark to lose trust in the data.

The 'WORK HISTORY SALARY UNIT' column had values of hourly, monthly, yearly and even other, which made it challenging to do any reliable analysis around the reported salaries. There were also some participants whose salary after going through the Occupational Skills training program was less that what they were making before. This again was interesting and confirmed some of the suspicions Hailey had about the training programs when she first spoke to me about them. I made note of all of these anomalies in the data as questions to ask Tara when I interviewed her next.

6.3.2.2 The Contextual Interview

At the start of the interview, Tara emphasized that she is primarily in charge of the data and didn't know as much about the program and its specific requirements. So while she could tell me where the data came from, she wasn't as sure about the programmatic rules about why certain data were collected on some participants and not others.

My conversation with Tara helped me better understand the structure of the data and clarify several of the questions I had. The first was the number of blanks entries in the dataset. Tara explained that the blank entries could have been a result of several issues. Some data might have been lost during the transition from the DoL database to the new GS database (which officially changed starting July 1st 2013). Others might have resulted from the new fields that GS introduced that could not be back filled for older records.

Others might have been a result of employees just not capturing the data at the time of entry.

I asked her about the OJT (On-the-job training) column, which seemed to have a lot of blanks in them. I pointed out one record of a female adult, whose work history suggested she was a housekeeper who made $5 an hour, but her OJIT column was blank. Tara said this meant the person did not receive on-the-job training, but might have received some other service, like job search assistance, counselling, other occupational skills training. This data was not in the spreadsheet Tara sent me because I didn't ask for it. Tara agreed to send me this additional data when I insisted that this would add a lot to the story we were trying to convey. Tara went on to explain that "*there is all sorts of data on these participants*", so every time she gets a request, she asks a lot of questions to ensure she understands what people want. In my case, she wasn't sure if I'd need all of the other columns which is why she did not include the columns about additional services in the dataset she sent me. Tara was willing to share any data that I needed as long as it did not contain any personally identifiable information about the participant. All we had to do was agree on the kind of data I was looking for, so Tara could write the appropriate SQL query to generate it.

I also brought up the 'ADDRESS' column, which I needed to determine which NPU each of the participants belonged to. Visualizing data based on their NPU's was one of the main features of the CWK dashboard and something all community leaders had come to expect . But for Tara, including the address column was not as straightforward because one might be able to trace the identity of the participant by combining it with other available data. The WIOA system collects three different addresses on each of the

participants, their current address, their mailing address (which could be different), and the address at the data of application. The program uses this last address to determine the participants eligibility and assign them to specific area-based training programs. Adding the address columns to the dataset required me to have a data sharing agreement in place with TCSG. This agreement would specify exactly what data we need and what we intend to do with it. Tara suggested I contact "the data guy at the State level", who sits in the department that oversees the WIOA program. My first few emails to him did not receive a response. While I waited to hear back, I explored other ways I might be able to classify the data I had (which included zip codes) into their respective NPU's.

6.3.2.3 Aggregating Records by NPU's

Determining the NPU from the address is a non-trivial task and requires advanced programming skills. When building the Public Safety module, our DSSG team wrote custom code to compare the address field of each record with the NPU boundary shape file. Not having access to the address field in the workforce training data, I contacted other experts to ask if there were alternate ways to classify data into their respective NPU's. Neighborhood Nexus (NN) does this will all of the data they release, so they were my first point of contact.

I spoke with Aaron who used to be a full-time employee in Emory's now-defunct Center for Community Partnerships, which was a core partner with NN. His work used to be part of Emory's contribution to NN. Now that Emory is no longer a core partner, Aaron works as a consultant to NN and specializes in reformatting the American Community Survey data that the Census Bureau collects.

Aaron explained to me the process he follows to make the annual ACS data usable for NN and other communities. The annual ACS release comprises over 54,000 fields, which is about 1.3 gigabytes of raw text. These fields only report counts and not the percentages based on those counts that one is used to consuming with Census data. For instance, Table B03002 from the 2017 5-year release of the American Community Survey reports a total population for the City of Atlanta of 465,230 and a Hispanic or Latino population of 21,450 but does not tell you that this 21,450 represents 4.6% of the total population. While there are some exceptions to this where some of the more detailed tables like the "demographic profile" tables do report percentages, percentages are not included as a general rule.

As anyone who has tried to navigate the Census data can attest, finding what you need within all this data can be daunting. The Census tool, American FactFinder, can be very useful if you want to get data from a table or two for a limited number of geographies, say to compare the data for the Census tract where you live to the county and/or state as a whole. But it is not as useful if one wants to look at the distribution of, say, the percentage of people in poverty across the region or explore the relationship between percent with a high school degree and percent in poverty. The main service that Aaron offers is to winnow down the available fields to a more manageable number, calculate relevant percentages, and impose a simpler set of categories to make the data more easily consumable. This data is then made available for standard Census geographies such as cities, counties, tracts, and state or federal legislative districts. These Census block groups or tracts (or groups thereof) generally serve as poor proxies for NPU's and City Council

districts. There are a number of possible strategies available to estimate these areas, each with its strengths and weaknesses.

The solution Aaron has settled upon is a population-weighted estimate based on Census block data. Census blocks are the "fundamental" unit of Census geography, they are the smallest available geographic unit from the Census Bureau and population counts are available at the block level through the Decennial Census. Aaron first uses Census blocks to build an area describing the custom geographies as closely as possible. Next, he uses the decennial population counts at the block level to determine the proportion of each block group's (or, if the data in question aren't available at the block group, tract's) population found inside each census block. That weighting factor is then multiplied by the variable of interest and then summed up to get estimates for the custom geography. A similar method is used to estimate the margin of error for each estimate. This margin of error for each block group gets inflated when split up into smaller pieces and then deflated again when those pieces get merged together.

All of this work is done through a series of scripts written in Stata (a statistical software for data science) that read the original text files, calculate percentages, and parse the data into something that can be fed into the WEAVE tool that NN use. These scripts get significantly modified annually when the Census Bureau announces the various changes that it is making to the following year's data release. Apart from matching the scripts to the latest data release, the process of repackaging the Census data has matured to a point where it is just a matter of downloading the data, running the script, and then checking that everything ran correctly.

Such a scripting process would be an overkill for the few thousand rows of workforce training data I had. Moreover, the only geographic information I had was the zip code, which was insufficient as a zip code could belong to more than one NPU.

I waited for a few weeks to hear back about the data sharing agreement and then looped in others I had met at TCSG, the Dean, along with Georgia Tech Legal, who I was told help out with such paperwork. I sent several emails and followed up for weeks but never heard back from any of them. I was running out of time and had to start building the visualizations so the new workforce module could be shared with the community leaders and become part of the data dashboard. I eventually decided to drop the address data request and make do with the expanded set of services data that Tara promised to share with me after the interview.

6.3.2.4 Creating Comparative Visualizations

After receiving the first sample dataset from Tara, Mark and I decided to explore it on our own and then meet up to exchange notes on what we had learned about it. Mark works as a consultant and uses tools like Power BI as part of his day job, which meant using Data Studio was not a huge leap for him. Being a first-time user with Data Studio, he occasionally asked me for some help on how to add and combine datasets in Data Studio. For instance, one email he sent me read

> *Working with GDS is proving a bit frustrating. One limitation is that we cannot use multiple data sources per report. For example, I have two sheets in my data set and want to use one sheet for two charts and the other sheet for the third chart but GDS does not appear to allow that. So it is hard to show multiple views of the data in one report ... especially since you cannot pre-save filters on the data.*

Any thoughts or suggestions you can provide would be much appreciated.

I replied with the following text that included a workaround and clarification on how I was using multiple sheets to generate a report.

GDS actually does allow you to add two datasets to a report. I just tried doing it in the report that I've shared with you. You should be able to see two pie charts in there, one which accesses data from the adult participants data source and the other from the youth participants one.

I was having trouble when both data sources were present as different sheets in the same excel file. I think GDS was getting confused and using just one of the sheets in that case. I had to separate out the sheets into individual data sources to get the report to work properly.

Hope this helps. Let me know if you continue to have trouble.

We continued our collaboration via email and sharing versions of reports with each other. When analyzing the workforce data, Mark was initially interested in knowing how many participants who joined the program ended up completing it and how many dropped out. He created the visualization below in Data Studio by counting the number of records in the 'COMPLETION STATUS' column.

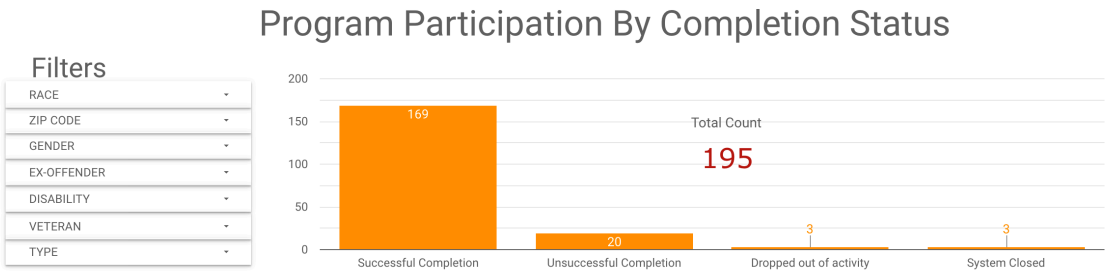


Figure 27: One of the Data Studio visualizations done by Mark, in which he counts the number of participants who completed the workforce training program.

The anomaly we noticed about wages before and after completing the Occupational Skills training program prompted Mark to compare the salaries of all the participants. The visualization below charts the 'HOURLY WAGE' after completion of the OST program with 'WORK HISTORY SALARY' for each of the occupations listed in the dataset.

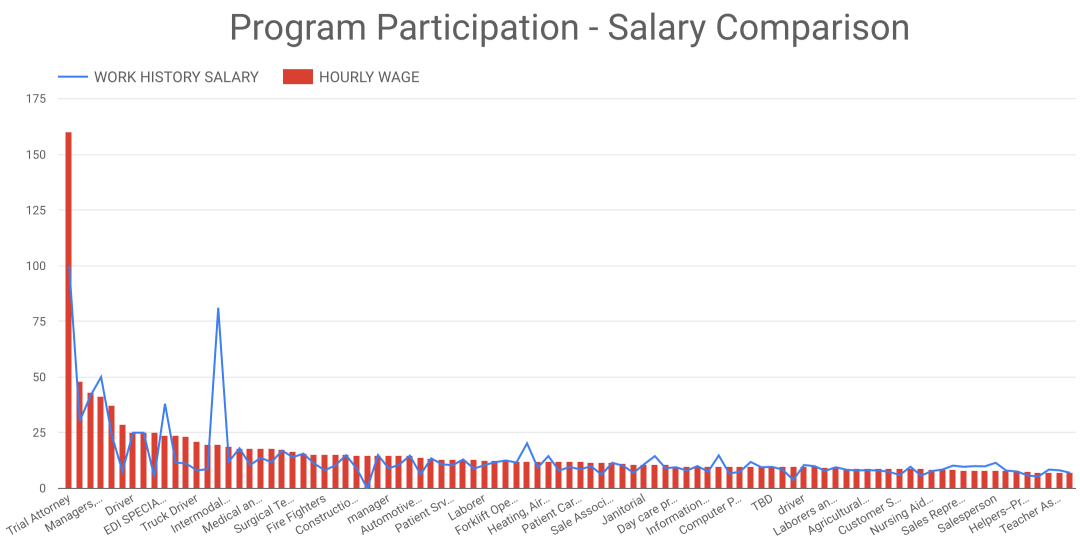


Figure 28: In this visualization, Mark compares to columns, 'Work History Salary' and 'Hourly Wage' of participants in the workforce training programs.

The visualizations that I created focussed on offering an overview of the demographic breakup of the participants who were participated in the workforce training program.

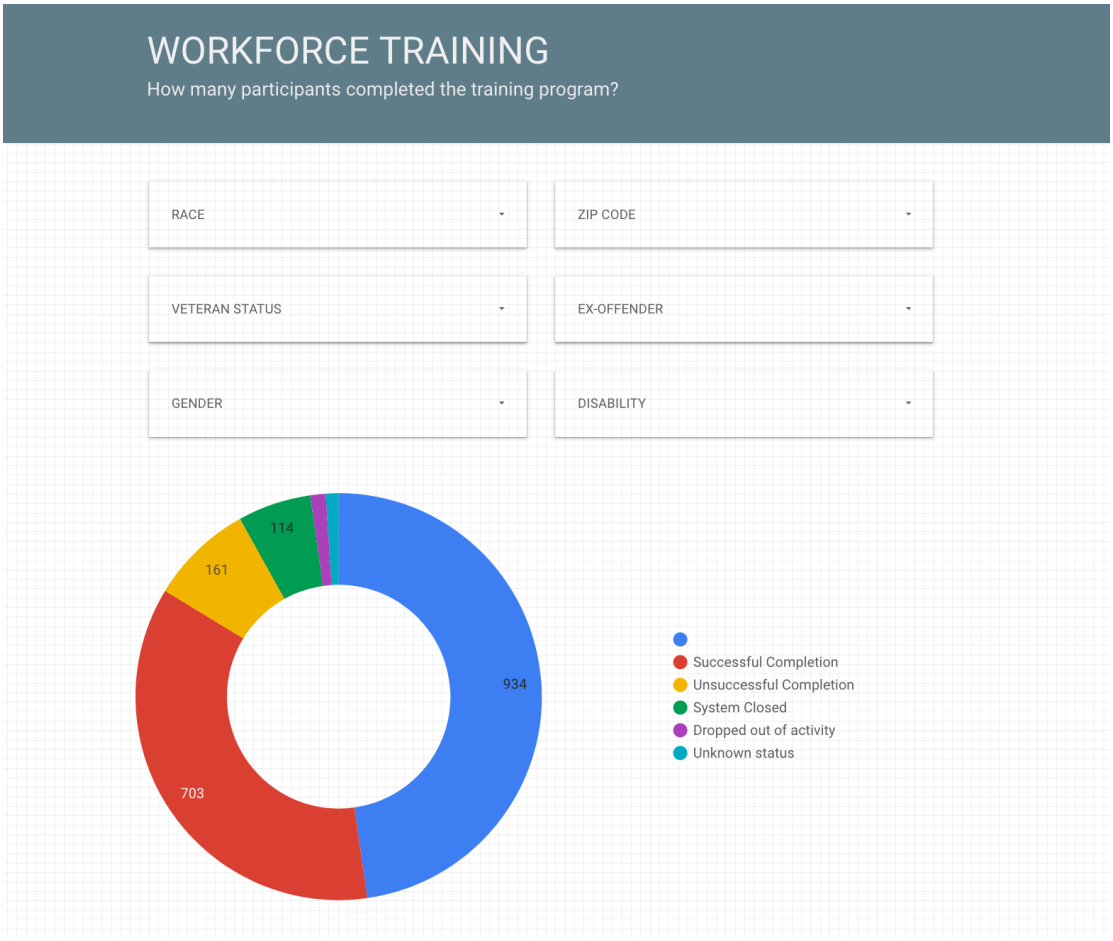


Figure 29: The Data Studio visualization I created, in which I compare program completion status by demographics.

6.3.3 Continued Participatory Infrastructuring with the Workforce Data

I continued exchanging emails with Mark over the next few months as we explored the data in more detail and thought about making it available on the dashboard as a module by itself. Mark was interested in sharing the insights from this dataset at one of his church meetings. I agreed to arrange for a projector and a screen in case he needed it for the presentation.

As we were working on the workforce dataset, Mark got interested in exploring data about the impact of library closures on the community. This library renovation project

was part of Fulton County's $275 million investment to overhaul 24 libraries, renovate 2 and open 8 new ones. Libraries across the county were being closed in phases in order to complete the renovations, which made Mark curious about the impact these closures were having on their communities. To learn more about this, Mark downloaded monthly circulation reports from the county library's website, which had data about book circulations, patronage and computer usage for each of the library branches. Since these reports were in PDF formats, Mark had to write a script that would extract the code from the PDF tables and insert them into Power BI. Once he had some clean datasets to work with, he emailed them to me so we could upload them to Google Drive explore them together in Data Studio. I also emailed Fulton County library's division manager and technical services administrator to clarify some of the doubts that came up during our analysis of the data.

Work on these datasets was progressing smoothly, when on July 20th, a few days before one of our planned meetings, Mark emailed me to say that he had met with an accident while riding his bicycle and his laptop was damaged beyond repair. He did have a laptop from work, but he wasn't too comfortable using it for his personal projects. This meant that Mark could not collaborate with me until he bought a replacement laptop, which he said would take a few months at least.

I continued working with the new data that Tara shard with me. This dataset included all the additional services that participants availed of through the program. For instance, it would include participants who availed not only occupational skills training but also orientation, job search/placement service, transportation assistance, individual

counselling etc. The dataset for the Adult Activity enrolments had 18001 records and the one for Youth Activity enrolments had 19,906. Both datasets had the following columns.

Table 9: Column names available in the activity enrolments dataset.

STATE ID	ZIP CODE	LWDA	WIOA APPLICATION ID	Activity Code
ACTIVITY DESCRIPTION	CUSTOMER GROUP	PROVIDER	TRAINING O*NET CODE/TITLE	ACTUAL BEGIN DATE
PROJECTED END DATE	ACTUAL END DATE	TRNG COMPLETION STATUS		

Tara also shared demographic data that was collected about the adults and youth who participated in the different programs. There were 16735 rows in the Adults dataset and 4928 in the youth dataset. Both datasets had the following columns.

Table 10: Column names in the demographic data about participants.

STATE ID	ZIP CODE	VETERAN STATUS	RACE	GENDER
AGE AT WIOA PARTICIPATION	HISPANIC	SCHOOL STATUS AT APPLICATION	BASIC SKILLS DEF AT APP 0=NO 1 =YES	ENGLISH LANGUAGE LEARNER 0=NO 1=YES
EX-OFFENDER	YOUTH IN OR AGED OUT OF FOSTER CARE	DISABILITY	WORK HISTORY JOB TITLE	WORK HISTORY OCCUPATION
WORK HISTORY START DATE	WORK HISTORY END DATE	WORK HISTORY SALARY	WORK HISTORY SALARY UNIT	LWDA
WIOA APPLICATION ID	PARTICIPATION DATE	PARTICIPATION FUNDING	CASE CLOSURE DATE	CREDENTIAL RECEIVED
DATE CREDENTIAL RECEIVED	JOB TITLE	OCCUPATION	HOURS PER WEEK	HOURLY WAGE
QTRTOTAL	START DATE	TRNG RELATED EMP	EXIT DATE	

Additionally, I also received one excel sheet that had all the activity codes and their descriptions listed. There were 103 different activities listed as rows in the dataset with the following columns.

Table 11: Column names in the activity codes dataset.

CUSTOMER GROUP	ACTIVITY CODE	SERVICE DEFINITIONS	SERVICE TYPE DESCRIPTION	TRIGGERS A NON-YOUTH PARTICIPATION
TRIGGERS A YOUTH PARTICIPATION	WILL NOT EXTEND A PARTICIPATION			

I created a new folder for all of this data in Google Drive. I saved the individual files as Google Sheets instead of Excel sheets as this makes importing them into DataStudio a lot easier. I combined these data in DataStudio to create the following visuals. An interactive version of these visuals is available on the dashboard website at http://dashboard.communitieswhoknow.com/Home/Workforce2.html. I describe this module in the next section and also point to specific ways in which it aligned with the data justice requirements and the long-term orientation of PD.

6.4 The Workforce Development Module

The process of reimagining the Westside community's data infrastructure and creating the workforce module required that I take a participatory approach that was oriented towards issues of long-term infrastructuring issues within PD. The new PD guidelines suggested by Bødker & Kyng (2018) that I summarized at the start of this chapter brought attention to the dual role of the researcher who is simultaneously involved in doing research and activism. New PD is action research that seeks out partnerships to create long term and scalable impact within a technology context. The emphasis is on creating 'working prototypes' with 'high technological ambitions' in order to influence change and impact that is sustainable. Similar concerns about sustainability and change

that is driven by the community feature in the universal values of design justice and data justice (Costanza-Chock, 2018; D'Ignazio & Klein, 2020; Dombrowski et al., 2016; L. Taylor, 2017), as well as the localized values of justice I developed in the previous chapter. Such an orientation towards justice was evident in my artful integration (L. Suchman, 2002) of the different Google platforms with the data ambassadors program, which led to the creation of the workforce development module.

The landing page of the Workforce Development module opens with an introduction in which I *disclose the context within which the data were created*. I do this using the Local Workforce Development Area map to explain where the data came from, how the different workforce service centers are structured and my role as a PhD student in collecting the data.

As the person who created these visuals in Data Studio, I currently have admin rights to this page. This means that I am the owner and control who else has access to it. Anybody who is interested in interacting with the visuals on this page can do so without having to download or install any additional software. This module was created with the Westside community leaders and ambassadors like Mark as its target group of users. These community leaders (CL's) can interact with the visual by changing any of the drop-down menus for race, gender, disability, ex-offender, zipcode, and training provider name. Changing the selections in any of these drop-down menus updates the data in all of the visuals and tables.

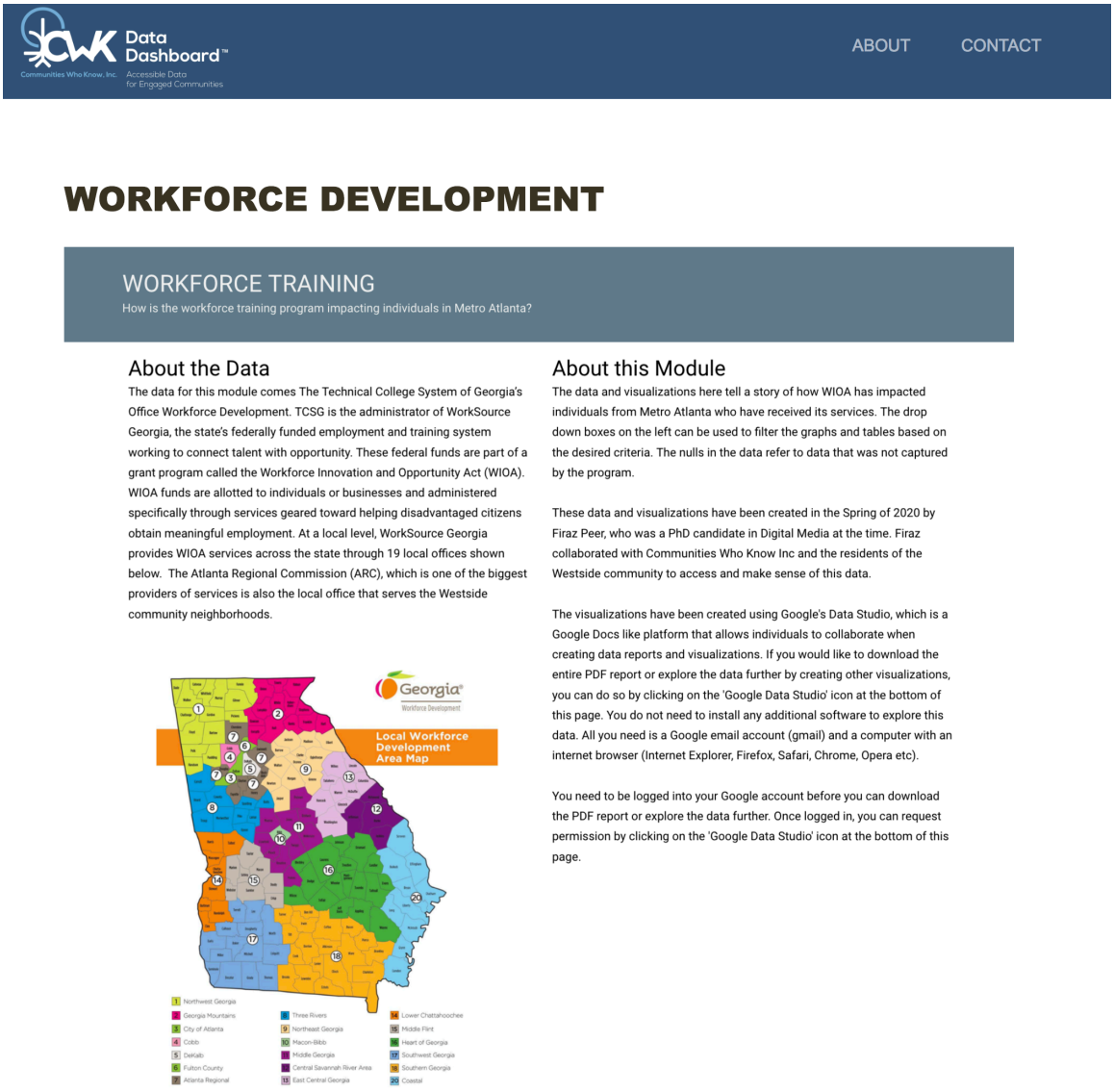


Figure 30: Interactive Workforce Training module on the data dashboard.

In order to change any of the visuals or see the data that was used to create them, CL's will need to open this visual in Data Studio and request 'edit' permissions. They can do this by clicking on the 'Google Data Studio' text at the bottom right of the visual. Once the CL's click on this text, the visual opens up in a new browser window, provided the user has 'edit' permissions. If they do not have the necessary permissions, an email is

sent to the admin (me in this case) asking for access. When requesting permission, CL's can see who the owner/admin of the visual is. The owner/admin then has the option to communicate with this CL and give them the permissions necessary to access and edit the visuals.

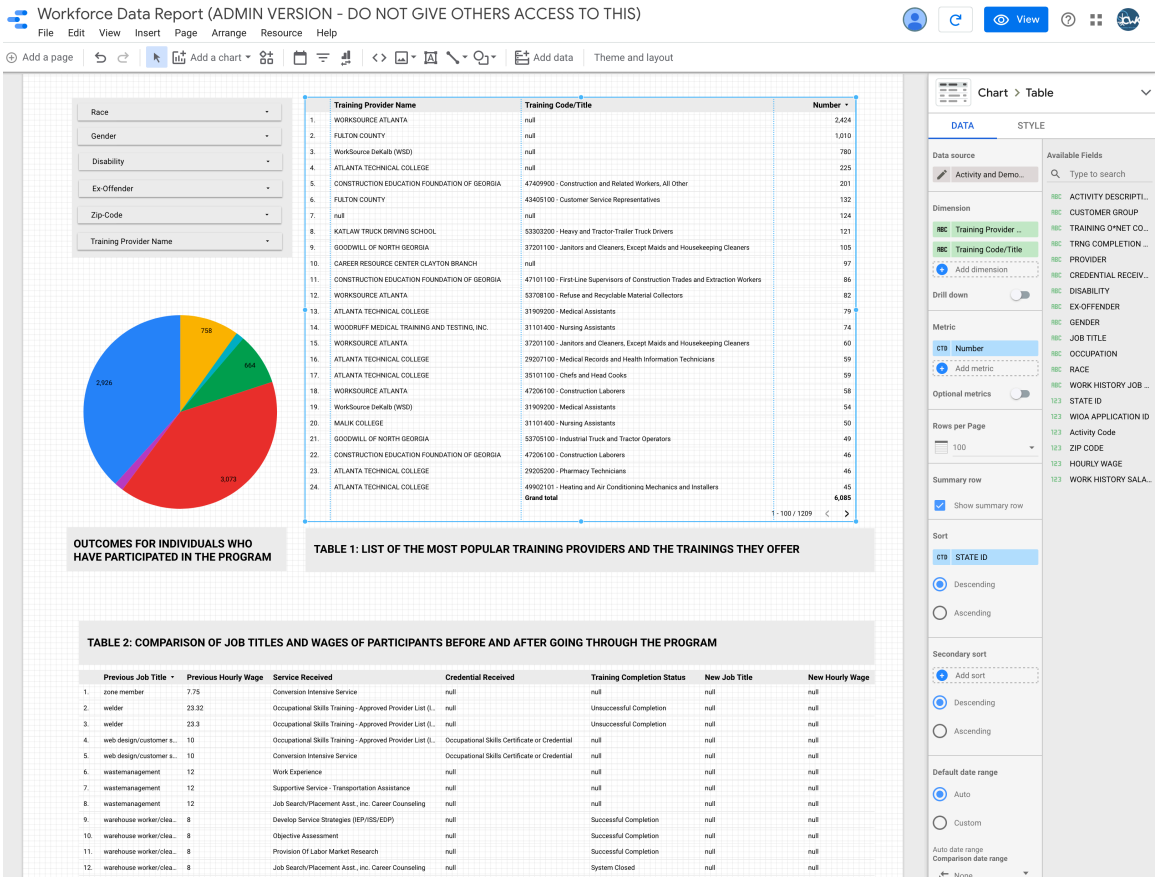


Figure 31: Interface seen by those who have 'Edit' permission to the data.

Since I worked with Mark to analyse the workforce training data and create this module, he is currently the only other person who has edit access to this module. Our partnership in creating this module helped reveal how data infrastructure literacy could be supported by such a platform. The platform supported a plurality of perspectives, as Mark created his own set of visuals with the data, which represented his perspective and his

interests in analysing the workforce dataset. Being the admin, I was *responsible* for ensuring that the data was free of discrepancies and up to date. I did this by contacting the owner of the data, who in this case was TCSG. Lines of ownership were thus clearly established as we know who was *accountable* for the data. As I described in the previous section, Mark contacted me whenever there was a discrepancy with the data or there was something specific with the Data Studio platform that did not make sense to him. He was planning on using the insights that came out of our exercise in some of his own presentations to his church. Mark was thus *empowered* to use this data towards his own advocacy purposes.

Maintenance and sustainability of the data are tied directly to Google's platform and the ability of the community leaders in managing it. The Workforce Module will be available as long as Google's servers are running and the Data Studio platform exists. The probability of Google's servers failing or the company ceasing to exist is much lower compared to the probability of an organization like the CWK ceasing to exist. But Google, like many other software companies have done in the past, can discontinue any of their products, which could in turn impact the alignment of these platforms as well. Mark or other CL's who wish to maintain the module, will need to become familiar with Google's platforms and understand how they are interconnected. These platforms can be replaced or augmented in the future depending on how the data needs of the community leaders change.

Even though the learning curve of setting up and maintaining a module like this in Data Studio is not as steep as doing it in Tableau or D3, it still requires a substantial amount of technological fluency (Disalvo & Lukens, 2009; Resnick et al., 1998), along

with data infrastructure literacy, and time that might be beyond what CL's can devote to such a project. Although desirable and asked for by the community leaders, such a literacy was at odds with the reality within which they operated. Put differently, the desired value of *supporting data infrastructure literacy* needed to be negotiated with the reality of the constraints in which the community leaders operated. Only one of the four data ambassadors who initially expressed interest in the program were able to follow through. This could be attributed to several reasons, which could include their lack of time, technological/data literacy skills, or just a lack of interest. In the end, it was only Mark, who had the time, interest, digital and data literacy skills needed to work with the data who fully participate in the program.

Apart from Mark, it was data intermediaries like Tara and Aaron who played a crucial role in enabling the workforce training module during the data ambassadors program. I discuss these data intermediaries next and conclude by discussing their relevance to the human infrastructure of the Westside's data infrastructure.

6.5 The Role of Data Intermediaries

Sawicki & Craig (1996) have described the different ways in which intermediaries serve community groups. These services fall along a continuum, from providing data, to converting this data to information, and finally converting this information into policies that can inform actual programs. These intermediaries are usually housed within large institutions like universities, public libraries, government offices and nonprofit groups. Librarians as intermediaries (Vitak et al., 2018), activists (Baack, 2015), journalists (Schrock, 2015), and hackers (Kubitschko, 2015) are just some of the data intermediaries that have benefitted from scholarly attention. In serving minoritized groups, these

intermediaries encounter many barriers like the lack of funding, technical skills, datasets about specific local concerns. González-Zapata & Heeks (2015) have also analyzed the barriers that developing countries face when making Open Government Data (OGD) usable by its citizens. These barriers, which include data absence, lack of data provision, lack of data quality and the digital divide increase the gap in the open data ecosystem between the public sector entities who produce the data and the publics that consume it. OGD Intermediaries help overcome these barriers by working as *demanders, producers, validators, developers* and *communicators* who convert data into information and knowledge that can be acted upon.

Sein & Furuholt (2012) offer vignettes of the work intermediaries do in offering basic Internet and government services in developing countries. They identified *Individual Entrepreneurial Intermediaries* as those who have setup businesses to offer basic services in exchange for a fee. These intermediaries also use their skills and resources to build a loyal customer base in order to sustain their business. They describe *Organizational Intermediaries* within the government or private sectors who are involved in determining policies and procedures that govern the distribution of services. There are also those who play a *multi-intermediary* role. These could be individual or organizational entities that offer services, consulting, training and even influence policy through advocacy efforts.

The sustainability of these individual and organizational intermediaries in developing countries has also been a subject of much research in ICTD. Ali & Bailur (2007; Kumar & Best (2006); Sein et al. (2008) build on their case studies in India, Bangladesh and Saudi Arabia, along with existing scholarship from ICTD to put forth five principles of sustainable projects financial/economic, cultural/social, technological,

political/institutional and environmental. Frequently, it is the inability to foresee such issues that causes most ICTD projects to fail. Since it is almost impossible to anticipate all of these issues with any accuracy, Ciborra (1992) has argued that such unexpected issues should be welcomed and incorporated into the project lifecycle. I also practiced an infrastructural version of bricolage that Dean Royster followed as she brought the dashboard together.

Chapters 5.3 and 6.1.1 described the kinds of hyper-local data that the community leaders needed in order to support their advocacy efforts. In some instances, such data had never been collected because nobody had tried to do so. There were also instances where data had been collected, but organizations lacked the time and resources to analyze and draw insights from them. When data were available, they were sometimes not the latest, were in formats that could not be easily translated or were privately held by companies which made it hard for community leaders to access them. Overcoming these barriers is beyond the capability of any single individual or organization and requires the co-operation of key elements of the infrastructure within which data are created and used. During my engagement through the data ambassadors program, there was just one instance, as with the workforce training data where the entire infrastructure came together to find the data we needed in the timeframe and format that made it usable for our needs.

Tara, the expert who gave me access to the workforce training data had access to all kinds of data about the program, but never distributed it because she wasn't aware of any demonstrated need to do so. When I told her about the dashboard I was working on, she was happy to share the data I needed as long as it did not violate any privacy concerns.

The process of finding the data, analyzing it and creating the workforce training module draws attention to the multiple data intermediaries that are involved in such participatory infrastructuring of data infrastructures. Data Intermediaries included individuals at organizations like TCSG, Fulton County, Neighborhood Nexus and Atlanta Regional Commission, who had roles to play in bringing the workforce training module together. These *Data Intermediaries* form the final piece of the human infrastructure of data infrastructure that I have been building through Chapters 4 & 5. I bring all of them together to build my recommendations for infrastructuring a community's civic data infrastructure in the next chapter.

Table 12: Elements of the human infrastructure (continued)

Infrastructural Elements	Humans that Put the Infrastructural Elements Together	Type of Human Infrastructure
Data Intermediaries like managers, database developers, programmers, statisticians etc.	Individuals at public/private/non-profit organizations, universities etc	Data Intermediaries

The participatory infrastructuring process I followed to counter the WCA dashboard and rethink the Westside community's data infrastructure towards more socially just outcomes serves as a case study for infrastructuring with social justice concerns in mind. This case study brings together scholarship social justice scholarship in HCI (Costanza-Chock, 2020; D'Ignazio & Klein, 2020; Dombrowski et al., 2016; Fox et al., 2016) with that on community based infrastructuring (Agid, 2016; Asad et al., 2017; C. Le Dantec & DiSalvo, 2013; Prost et al., 2019; Vasilis Vlachokyriakos et al., 2018b) to offer insight

into what it takes to infrastructure for civic data within resource constrained environments.

I initiated this process of infrastructuring by identifying a localized set of data justice values needs through the workshops. These values and needs served as hypotheses during the infrastructuring of the Westside's civic data infrastructure. In negotiating these values through the infrastructuring process, I realized how a value like supporting data infrastructure literacy was in conflict with the resource and time constrained reality in which the community leaders operated. This furthers the argument made by Grönvall et al. (2016) and Parvin et al, publishing as JafariNaimi et al., (2015) that values are situated/dynamic, which is why they cannot be *applied* to design situations. The infrastructuring process allows designers to negotiate which values might best serve the situation at hand.

My infrastructuring process involved integrating Google's Data Studio, Drive and Groups platforms together with a data ambassadors program which brought community leaders and data intermediaries across Atlanta together. I wore multiple hats through this infrastructuring process. At times I worked as a facilitator, trying to organize the data ambassadors program, which involved bringing community leaders together finding data that was most relevant to their needs. I was the expert when it came to interpreting the values that were discussed in the workshops and selecting relevant data platforms that could satisfy them. In executing these roles, I remained conscious about not taking the lead on creating data stories myself but allowing the participants to express their perspectives through the data.

As a fully functioning prototype, this infrastructure demonstrated the potential to influence change that could largely be community led and sustained, provided the community leaders had the time, interest, digital/data infrastructure literacy skills to do so. *Data Intermediaries* emerged as another key element of the data infrastructure who ensure that the data are available in specific formats and usable by others. I bring these data intermediaries and other human elements of the infrastructure I identified in previously together into a final set of contributions in the next chapter.

CHAPTER 7. ALL DATA ARE HUMAN

This dissertation is grounded in the field of HCI. More specifically, I draw from subfields within HCI, which are concerned with redesigning the infrastructural elements that underlie our digital civic information and communication technologies from a social justice perspective. I initiated this dissertation with the question, *What are the local values and infrastructural arrangements that are required to build, use and maintain equitable data infrastructures that enable communities to benefit from the publicizing of data through dashboards?* The different phases of my fieldwork allowed me to unpack the local values and infrastructural arrangements that it took to build, use and maintain data infrastructures from the perspective of the community leaders I was working with. My research site was restricted to the Westside communities in Atlanta, which I scaled by relying on participatory methods of studying infrastructures.

I started off by considering the Westside Communities Alliance Data Dashboard as an assemblage and performing an ethnography that allowed me to unpack the praxis, politics and infrastructural mis/alignments at play when building such civic data infrastructures for marginalized communities. Observing the specific infrastructural breakdowns and the Dean's artful integration of her human, material and organizational resources in response, led me to identify infrastructural bricolage as one possible strategy to bring such a dashboard to fruition within resource constrained environments. But this artful integration, as far as I could tell, came about through a research agenda envisioned by the Dean and not through a participatory process that involved the community in its design and development. The dashboard was therefore unjust from the perspective of

participatory design and social justice scholarship in HCI. This prompted me to ask whether a data dashboard was the best way to bring about data equity and consider alternate infrastructural arrangements that might meet the Westside community's data equity needs. I hypothesized that an infrastructuring approach that was community led and prioritized local values and needs would have a better chance of realizing data equity and coping with infrastructural breakdowns in the future. I organized data literacy workshops in order to surface these local values and needs, which guided the manner in which I reimagined the Westside community's data infrastructure. Taking a participatory approach to infrastructuring, I reimagined the Westside community's data infrastructure by integrating off the shelf platforms and working with community leaders to build their capacity to use and maintain these platforms. Taken together, this infrastructural bricolage of human, material and organizational resources, participatory infrastructuring of data literacy skills and infrastructures offers one answer to my proposed research question. I expand on the specifics of these contributions and their alignment with the study of civic data infrastructures in HCI through the rest of this chapter.

7.1 The Human Infrastructure of Civic Data Infrastructures

Building on Bowker & Star's (1999) concept of infrastructure, Lee et al. (2006, p. 484) define human infrastructure as the "arrangements of organizations and actors that must be brought into alignment in order for work to be accomplished". The definition implies that the nature of the work is distributed and involves multiple actors and networks working in alignment with each other. Sambasivan & Smyth (2010) propose a hubs and spokes model to visualize human infrastructures, which depicts strong ties and weak ties between its actors who control access to resources. These actors are also in charge of

building capacity to use the infrastructure and also restoring it after moments of breakdown or crisis. In this way, human infrastructures are 'multimorphous', in that they are messy, chaotic, and difficult to manage as they are constantly changing over time. These changing elements of the human infrastructure were brought into sharp focus through my three-step process of performing an infrastructural inversion of the dashboard, building the community leaders' data infrastructure literacy and reimagining the Westside community's data infrastructure. I draw attention to these elements of the human infrastructure next, while also reflecting on my changing role throughout this project and the empirical contributions that resulted from them.

I first engaged with the Westside data infrastructure as an ethnographer/developer, working in collaboration with the WCA to build the Public Safety module of the data dashboard. The ethnography of the dashboard revealed the infrastructural elements and alignments that were involved in bringing the dashboard to fruition. In unpacking these sociotechnical black boxes that made up the dashboard, I engaged with *Community Champions, Transient Workers, Institutions,* and *Organizations* that that were involved in infrastructuring the dashboard.

Community Champions are actors who have the power to organize resources around specific matters of concern and influence change. The Dean is an example of one such community champion. Her efforts in setting up the WCA and then the CWK as community organizations that served the community through several projects are evidence of such power and influence. Examples of other such actors would include mayors, elected officials, university presidents, researchers and other individuals across the city, who have the power and resources to influence change.

Transient Workers are the students, faculty, staff, researchers and other actors whose involvement within the data infrastructure was limited to specific time periods, modules or events. The GRA's who only worked on the dashboard for a semester or two, students who built specific modules through their participation in the Data Science for Social Good program and the visualization course are examples of such transient workers. The staff at the WCA as well as other faculty at Georgia Tech and neighboring institutions also had a role to play in the infrastructuring of the dashboard. As a design researcher studying the dashboard for a specified amount of time, my work during specific phases of the project could also be categorized as transient.

Institutional actors are the ones that establish the policies, protocols, funding agreements, tax laws, licenses and other rules that govern how other actors within the infrastructure interact with each other. These institutional actors can be city, state or federal agencies, university boards and others with similar kind of power and influence.

Through this process, I interacted primarily with the *Community Leaders* who were active in the Westside. These community leaders were critical in determining what it meant to do justice to data. They serve as gatekeepers and knowledge repositories for anyone who is seeking to engage the community through specific matters of concern. Many of the community leaders I spoke with are also involved in *organizations* that are aligned with specific matters of concern. These organizations give community leaders a formal structure to organize, fund and execute on their advocacy efforts. As observed with the WCA, the value and longevity of these organizations depends on the individuals who run them, the kind of work they do as well as the resources and influences they bring to the table.

One of the motivations for these community leaders to organize was to improve the public perception of their neighborhoods and present them favorably to those who are *External* to their neighborhoods, like researchers, funders, media outlets and the wider public. This was important, the community leaders claimed, because these external actors contributed to the public perception of the Westside neighborhoods. Researchers and media outlets used publicly available data to represent the Westside in a way that was different from the ground reality that the community was experiencing. Working as an infrastructural bricoleur, Dean Royster attempted to intervene in this situation through the artful integration of data through a variety of sociotechnical platforms. *My dissertation therefore offers evidence of how the concept of infrastructural bricolage, which has previously been discussed in the domain of information systems* (C. Ciborra, 2007)*, computer programming* (Turkle & Papert, n.d.) *and workplace practices* (Erickson & Sawyer, 2019) *can also be applied to the design of civic data dashboards and their underlying infrastructures. This would be useful for designers, researchers and practitioners who are looking to build similar data dashboards and infrastructures for their communities. Additionally, in using a data assemblage like the WCA dashboard as a window into the infrastructural inversion of the Westside community's data infrastructure, I help position the study of data dashboards within the wider field of information infrastructures.*

While such public availability of data via dashboards and digital media authoring tools are said to be democratizing the access, interpretation and use of data, they are doing so only within privileged communities of practice. The data literacy workshops I organized

was one attempt to bridge this data and digital divide between the privileged and marginalized communities.

When organizing the workshops, my role changed to that of an expert. I was the expert in the room who was in charge of ensuring that the data and activities I chose were relevant and interesting to the workshop participants. I structured the workshops so they served as boundary objects for both data infrastructure literacy and design. The workshops were punctuated by several moments of breakdowns which exposed the various infrastructural layers through which data came together and also pointed to unmet needs and values that one needed to consider in future iterations of the dashboard. Experiencing these breakdowns and seeing how they led to design insights led me to develop a set of workshop guidelines for data infrastructure literacy. The workshop guidelines ask that we

1. Invite participants for whom the data and the issues it represents are *matters of concern.*
2. Take a *critical pedagogical approach* that aims to contextualize the workshop activities with specific matters of concern that the participants identify with.
3. *Be open to an infrastructural imagination,* where participants reimagine the data infrastructure by critically analyzing its black boxes and barriers.
4. *Focus on moments of breakdown* and be open to discussing inaccuracies and errors in the data infrastructure.

Additionally, these workshops also revealed local values of data equity from the perspective of the community leaders who participated in the workshops. *These values are a response to LeDantec et al's. (2009) call for more prescriptive methods for the*

discovery of values in design. The needs and barriers to accessing, using and interpreting data that surfaced during the workshops served as building blocks for a localized values of data equity, which is how L.Taylor (2017) recommends we operationalize her data justice framework. These values of data equity were specific to the time and place I was designing in and ask that we

1. Support the community's data infrastructure literacy.
2. Empower the community towards action.
3. Foster accountability and responsibility in and through the data.
4. Disclose the context within which data are created.

I used these values as hypotheses (Parvin et al, publishing as JafariNaimi et al., 2015) in the subsequent reimagination of the Westside data infrastructure. The process involved taking a participatory infrastructuring approach to reimagine the Westside community's data infrastructure in a manner that prioritized the localized values of data equity I had discovered through the workshops. This participatory infrastructuring process was further guided by an orientation towards justice that is reflected in the recent HCI and PD literature. My role in this phase of the project was that of a facilitator and researcher rather than that of an expert or activist taking the lead on matters of concern. My role was limited to infrastructuring the various elements of the Westside community's civic data infrastructure in a sustainable and collaborative manner. This included working with many *Data Intermediaries* across the city, along with the other human infrastructural elements that I encountered previously.

Data intermediaries represent the actors who have the specialized knowledge and skills to create the data that others need. These data intermediaries interact with primary

data sources to reproduce data into formats that can be interpreted and used by other actors who lack such skills. I interacted with individuals in organizations like Neighborhood Nexus, Atlanta Regional Commission, Georgia's Office Workforce Development and Emory University to get access to data in formats that best served the Westside community's infrastructuring needs. These data intermediaries are typically more concerned with the quality of the data they put out and its usability than with the specific matters of concern that drive the community leaders. I summarize all these elements of the human infrastructure in the table below.

Table 13: The human infrastructure I engaged with in infrastructuring the Westside community's civic data infrastructure.

Infrastructural Elements	Humans that Put the Infrastructural Elements Together	Type of Human Infrastructure
Dashboard, funding, organizational structure, management	Dean	Community Champions
Individual dashboard modules, interactions with the community, software platforms	Graduate Research Assistants, Other Researchers, Faculty and Staff	Transient Workers
Communications network (newsletter, website, dashboard), programs, workshops, data/dashboard/infrastructure maintenance, Westside Resource Center	Westside Communities Alliance/ Communities Who Know Inc., Other Non-profit Organizations	Organizational
Data, policies, protocol, policies, funding, philanthropy, tax regimes, laws, licenses, jobs	Political and Educational Institutions. Schools, Colleges, City, County, State and Federal Agencies	Institutional
Advocacy efforts which include organizing meetings, programs, events, festivals, workshops etc	Community Leaders from different communities in the Westside.	Community Leaders

Media organizations, other researchers, communities.	Individuals at local and national news organizations and universities.	External
Data Intermediaries like managers, database developers, programmers, statisticians etc.	Individuals at public/private/non-profit organizations, universities etc	Data Intermediaries

In spite of taking a participatory approach to infrastructuring that prioritized the Westside community's data needs and values, the fact remains that only one community leader interacted fully with the data infrastructure I created. This community leader was digitally literate, had many of the foundational data literacy skills that I was hoping to build through the workshops and also had the interest to put in the time it took to analyze the data and create relevant stories. Ultimately, his participation was also curtailed by breakdowns that could not be preempted. The redesigned infrastructure failed to serve all the changing needs and constraints that defined the work of the community leaders.

This furthers the argument made by Grönvall et al. (2016) that values are dynamic and need to be negotiated throughout the design process. Values like supporting data infrastructure literacy were in conflict with resource and time constrained environment in which the community leaders were operating. My attempt at integrating this value into the reimagined Westside data infrastructure did not bring about significant engagement from the community leaders as they were unable to commit the time and resources that were required to access, interpret, and use the data.

The community leaders I worked with demanded data about hyper local issues that were, in many instances, not being collected at the state or national level. Visualizations and dashboards that convey data about census, public safety and the like are available in plenty. What communities need help with is to find, interpret and use data about issues

that are hyper local to their communities. As demonstrated with the data ambassadors program, such hyper local data can be found by mobilizing key elements within the data infrastructure. Finding data or collecting it, cleaning it, analysing it and using it requires expertise and time that not everyone have. Communities would be better served by identifying individuals with the right set of skills who can support such data advocacy work rather than investing in publicising data through dashboards and expecting engagement in return.

Building the community's data infrastructure literacy might make people more vigilant and aware of the inherent biases in data, but the act of working with data and creating stories requires a different set of skills, expertise and time, which not everyone can afford. In addition to the time it takes to develop these digital and data literacy skills, one also needs to consider the interest, time and effort it takes to find the required datasets, interpret them and use them in ways that best suit the community's matters of concern. This process is mired in layers of infrastructural alignments that are best navigated by individuals and organizations who possess the skills, interest, time, power and influence to do so. *Building the workforce training module offered evidence of the praxis and politics involved in the infrastructuring of civic data for social justice, which is a growing area of interest in the study of information infrastructures* (Agid, 2018; Halkola et al., 2015; Korn et al., 2019; Prost et al., 2019; Vasilis Vlachokyriakos et al., 2018b).

The design anthropological lens I used to unpack the Westside community's data infrastructure highlights the complicated reality of initiatives that hope to bring about data equity for resource constrained communities. On the one hand, it highlights the

injustice of building civic data dashboards in which the community does not actively participate. On the other, it highlights the complexities of participatory infrastructuring in such communities, which are socio-economically burdened and whose leaders are actively engaged in advocating for change. It highlights the need for the long term infrastructuring of civic data that not only involves an assemblage of technical platforms, but also a social network of individuals and organizations that can build, use and maintain a data infrastructure that is nimble enough to adapt to unexpected breakdowns and the changing sociotechnical context. Such a sociotechnical analysis of the data infrastructure, which includes a specification of the actors and networks, their values, power asymmetries and complexities involved in the infrastructuring of civic data is the novel contribution I make in this dissertation. Similar theories and frameworks for long term infrastructuring exist, but are mostly concerned with scientific research infrastructures (Bietz et al., 2010; Lee et al., 2006; Randall et al., 2015; Ribes & Finholt, 2009) and civic engagement (Balestrini et al., 2017; J. Carroll & Rosson, 2013; Gordon & Mihailidis, 2016; Rogers & Marshall, 2017). In making this research particular to civic data and their underlying infrastructures, I take the first step towards building a framework that can guide researchers and practitioners in building long term civic data infrastructures with their communities.

The approach to infrastructural bricolage that involved the artful integration of specific humans, materials, organizations, which was practiced by the Dean when building the dashboard and myself when reimagining the Westside's data infrastructure is one possible response to the more recent critiques of PD, which have asked that researchers deploy working prototypes in order to create sustainable, long term impact.

Infrastructural bricolage as discussed in this dissertation offers a model for the infrastructuring of working prototypes that can be updated and re-deployed based on the community's changing needs over an extended period of time. It is impossible to anticipate all possible infrastructural breakdowns and have a contingency plan in place, so the practice of infrastructural bricolage can help design researchers and practitioners adapt their infrastructure in the face of such inevitable breakdowns.

Observing the coming together of these elements of the human infrastructure in order to design, use and maintain the Westside community's data infrastructure through a localized conceptualization of values leads to one possible answer to the overarching research question driving this dissertation. Bringing these elements of the human infrastructure together and reflecting on how my role as a design researcher changed during the scope of this project, I argue that *all data are human, and the way we do justice to them is by identifying and building relationships between the human elements of the civic data infrastructures that we are trying to build.* This implies that we focus on identifying the human actors that are crucial to the civic data infrastructures we are trying to build, strengthen their working relationships and prioritize their values and needs by including them in our infrastructuring efforts. Specific technologies should be chosen and developed based on the needs, values, skills and relationships of these human actors rather than having them adapt to the manner in which technologies work. The diagrams below help develop a visual representation of the human infrastructure and their sociotechnical alignments that have described so far.

Figure 32: Developing a visual representation of the human elements of the Westside community's data infrastructure.

I start off by centering the community's *matters of concern*. This is similar to the diagram put forth by (Balestrini et al., 2017), in which they focus on orchestrating large scale citizen engagement around common matters of concern. Examples of some of these concerns in the Westside are summarized in Table 14 below. This list is not exhaustive and will vary with time and the sociopolitical context within which one is interacting with such communities.

Table 14: Examples of different matters of concern in the Westside.

Affordable housing in the Westside neighborhoods.
Students' performance in elementary schools
Orientation booklet for churches.
How do culture and arts impact youth and pride in the community?
Where can we put community gardens in the neighborhood?

How many youth service programs are running in the neighborhood?
Data to help the community newspaper reach a broader audience.
How are a neighborhood park and recreation facility being used?
Data about resident engagement with NPU's.
Organizations doing impactful work in the Westside
Impact of food insecurity on children's school performance
Respond to G-DOT's Renew Atlanta $400m shortfall
How are job training programs they stabilizing families?

After establishing common matters of concern (Balestrini et al., 2017) ask that we frame the issue from different perspectives before identifying viable solutions. These solutions, which could be tools, interactions, management protocols or learning structures need to be deployed, tested, iterated and improved upon in situ so we can identify sustainable ways in which the best practices can be scaled and shared. They also emphasize the centrality of humans in this entire process, which is critical to my representation as well.

My research has established the centrality of humans and their relationships when infrastructuring civic data infrastructures for communities. These humans are influenced by other individual, organizational and institutional actors who also have the power to shape agendas and outcomes. All of these infrastructural elements, which include with *Community Champions, Transient Workers, Institutions, Organizations, Community Leaders, External Actors* and *Data Intermediaries* are represented in the figure below. Each of the figures represent all of the actors placed around the center based on the power and influence they might have over specific matters of concern. Since power and

influence are dynamic and situational depending on the specific actor and matter of concern one is dealing with, I represent them using dotted orbital lines.

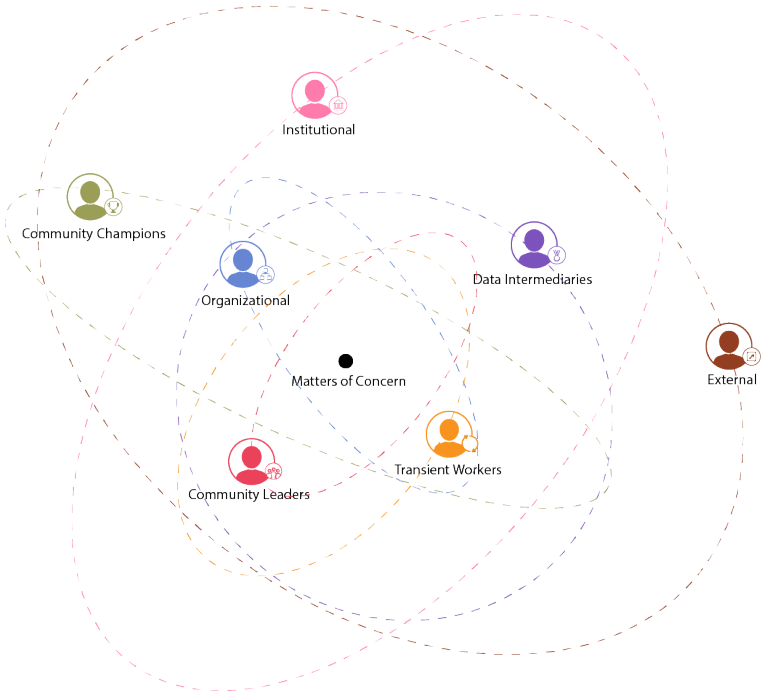


Figure 33: Doted orbital lines depicting the changing power and influence actors have over specific matters of concern.

The dotted orbits in Figure 33 represent the changing influence that the actors have when it comes to specific matters of concern. Community leaders and organizations for instance, are a lot closer to the center because they are a lot more involved in community concerns and have much more influence when compared with other actors. Other actors, like organizations, institutions, transient workers, data intermediaries, external actors and even community champions tend to rely on these community leaders and organizations to better understand local concerns and influence policies. It is therefore crucial that one interacts with the community leaders early on to understand their perspectives and priorities to design data infrastructures that are centered around them. The organizations

one interacts with will depend on the specific concerns and data one is working with. These interactions between the different actors are represented by the intersections between the orbits in Figure 33.

The figures below (34 – 42) are meant to add further clarity to how I developed Figure 33. Each figure below represents a specific actor I identified and their corresponding dynamic orbit of influence over the matter of concern. I developed the final representation by layering each of the actors and their corresponding orbits in the following figures on top of each other.

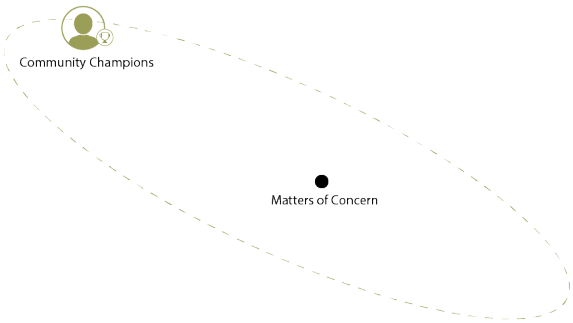


Figure 34: Community champions and the orbit representing their dynamic influence and distance from the matters of concern.

Figure 35: Community leaders and the orbit representing their dynamic influence and distance from the matters of concern.

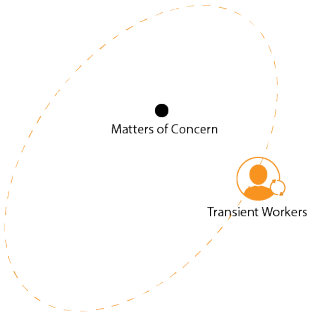


Figure 36: Transient workers and the orbit representing their dynamic influence and distance from the matters of concern.

Figure 37: Organizational actors and the orbit representing their dynamic influence and distance from the matters of concern.

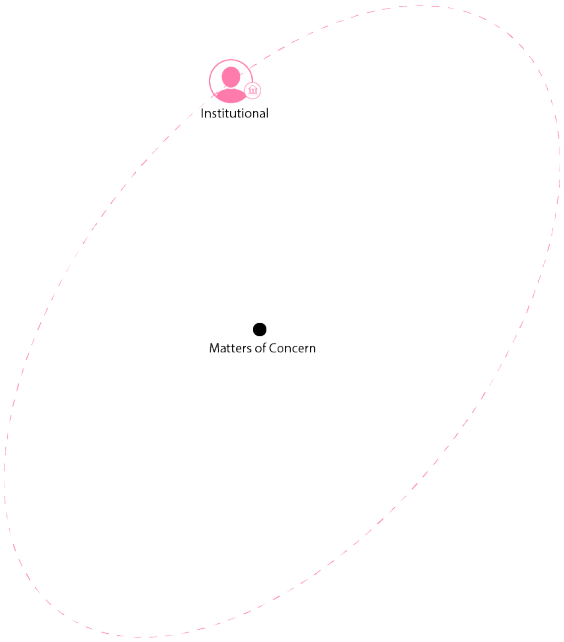


Figure 38: Institutional actors and the orbit representing their dynamic influence and distance from the matters of concern.

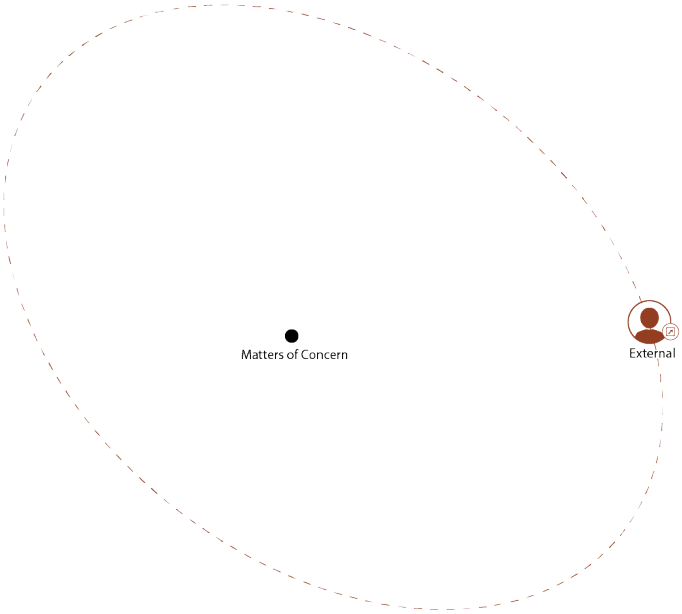


Figure 39: External actors and the orbit representing their dynamic influence and distance from the matters of concern.

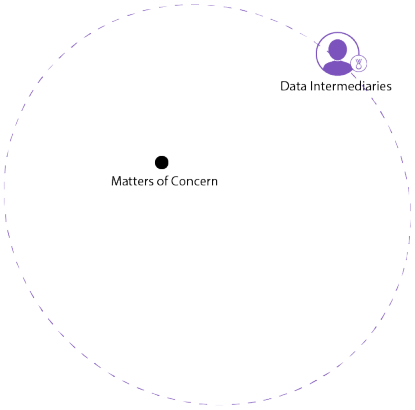


Figure 40: Data intermediaries and the orbit representing their dynamic influence and distance from the matters of concern.

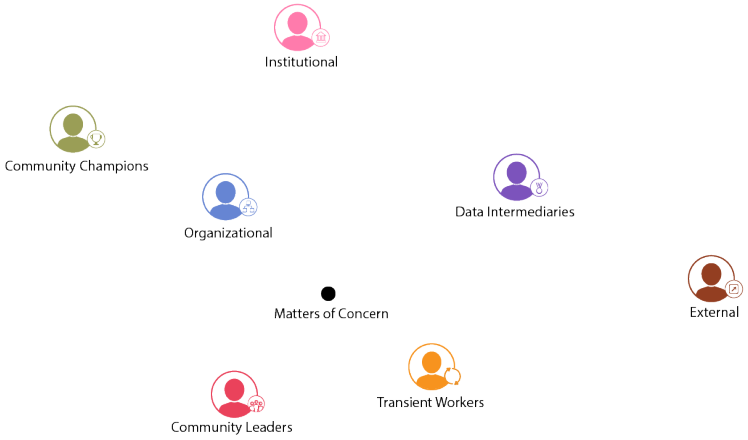


Figure 41: Doted orbital lines depicting the changing power and influence of the human infrastructure.

Based on the recommendation by Taylor (L. Taylor, 2017), I developed localized values of data justice that helped surface the specific functionings and capabilities that communities wish to prioritize. Such a list allowed me to produce values for data equity that are grounded in the local experience of the community leaders for whom the data matter the most. These values are represented in the final diagram below.

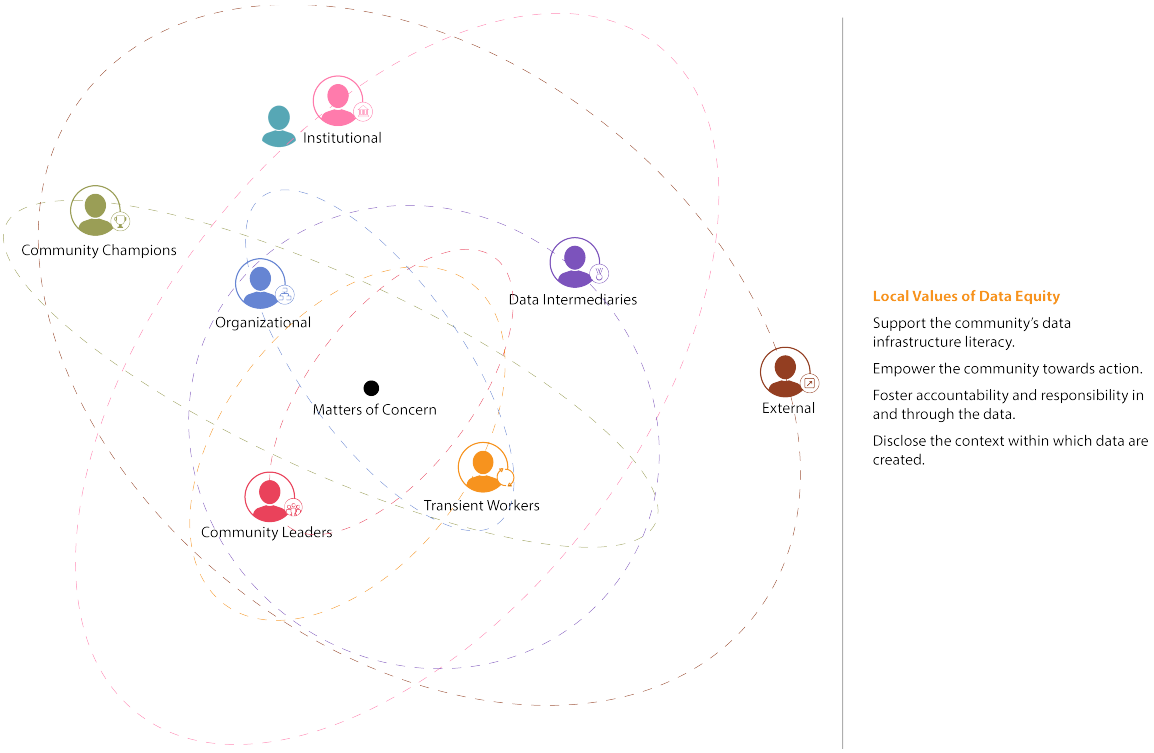


Figure 42: Diagram depicting all of the human infrastructure, their alignments and the local values of data equity involved in infrastructuring civic data with communities.

This diagram is but one representation of the capabilities, functionings, values, agencies, spheres of influence, matters of concern, praxis, politics and alignments of the human infrastructure that are involved in infrastructuring civic data for and with communities. These findings are intended to initiate discussion, thought and future work about what it takes to design just data infrastructures that are grounded in the local

experience of the communities one is infrastructuring for and with. My results are grounded in the specific experiences of a group of marginalized communities in Atlanta, which should be critiqued and expanded upon when being applied to other communities, whether marginalized or not. The specific elements of the human infrastructure as well as their alignments, influences, power asymmetries would vary based on the community one is infrastructuring for and with. The infrastructural elements and their alignments could change even in the same community, depending on which actors are involved in the design of the data infrastructure. I discuss these limitation and future implications of this work in the next chapter.

CHAPTER 8. CONCLUSION

This dissertation was motivated by the call to bridge the data divide that has been identified by scholars within Community Informatics, Communication and Critical Data Studies (Andrejevic, 2014; Boyd & Crawford, 2012; Gurstein, 2003). These scholars are concerned that the benefits that come from the democratization of data and analysis tools are not equally distributed in society. There is a gulf between the data have's, who have the knowledge, resources, skills needed to access, interpret and use data for advocacy and the data-have-not's who lack these privileges and are therefore being left behind in the ongoing data revolution. Data dashboards are tools that are being developed in the hope of bridging this data divide, by offering communities the data they need when advocating for change.

The WCA Data Dashboard was one such tool that was built for the Historic Westside communities in Atlanta. Initiated by the Dean of the Ivan Alan College of Liberal Arts at Georgia Institute of Technology, the dashboard was an output of a university-community partnership that the Dean hoped would forward her community-based research agenda, while benefitting the community as well as the university. The Westside community would benefit from the talent and resources available at the university, while the Dean and the university would have the opportunity to do good in their neighborhood.

The dashboard went through a multi-year development process in which it faced a number of setbacks and breakdowns. In performing an ethnography of the dashboard and its development process, I was able to identify the human, material and organizational arrangements that are required to build, use and maintain equitable data infrastructures

with and for marginalized communities. The conceptualization of equity came from local and contextual accounts of what it meant for community leaders to do justice to data.

The findings from my dissertation and my methodological approach have direct implications for the field of HCI. Each chapter offers specific contributions to critical discussions civic data infrastructures. In Chapter 4, I offered evidence of *infrastructural bricolage*, which has previously been discussed in the domain of information systems (C. Ciborra, 2007), computer programming (Turkle & Papert, n.d.) and workplace practices (Erickson & Sawyer, 2019) can also be applied to the design of data dashboards and their underlying infrastructures. This would be useful for designers, researchers and practitioners who are looking to build similar data dashboards and infrastructures for their communities. Additionally, in using a data assemblage like the WCA dashboard as a window into the infrastructural inversion of the Westside community's data infrastructure, I help position the study of data dashboards within the wider field of information infrastructures. Finally, the infrastructural inversion showed how building a data dashboard without the active participation of the community furthered many of the injustices that the dashboard was trying to combat in the first place. The socially good intention of making civic data freely available via a data dashboard ignored the distinction between data access, interpretation and use that has been discussed previously (J. Carroll et al., 2018; Gurstein, 2003, 2011; Puussaar et al., 2018). It is this observation that led me to propose a social justice-based approach to infrastructuring civic data that is grounded in the community's local data needs and values.

The data literacy workshops I organized served as boundary objects that helped specify what it takes to support data infrastructure literacy within community settings

while also revealing their local data needs and values. The workshop guidelines for data infrastructure literacy ask that we

1. Invite participants for whom the data and the issues it represents are *matters of concern.*
2. Take a *critical pedagogical approach* that aims to contextualize the workshop activities with specific matters of concern that the participants identify with.
3. *Be open to an infrastructural imagination,* where participants reimagine the data infrastructure by critically analyzing its black boxes and barriers.
4. *Focus on moments of breakdown* and be open to discussing inaccuracies and errors in the data infrastructure.

These guidelines offer one possible avenue for furthering data infrastructure literacy (J. Gray et al., 2018) within communities. Additionally, these guidelines serve as a response to Le Dantec et al's. (2009) call for more prescriptive methods for the discovery of local values in design. The needs and barriers to accessing, using and interpreting data that surfaced during the workshops served as building blocks for local values of data equity, which is how L. Taylor (2017) recommends we operationalize her data justice framework. When discussing their needs and aspirations with regards to civic data, the community leaders asked that it

1. Support the community's data infrastructure literacy.
2. Empower the community towards action.
3. Foster accountability and responsibility in and through the data.
4. Disclose the context within which data are created.

These values of data equity served as hypotheses (Parvin et al, publishing as JafariNaimi et al., 2015) during the subsequent reimagination of the Westside data infrastructure. I reimagined the Westside's civic data infrastructure through a participatory infrastructuring process in which community leaders took on the role of data ambassadors and created data stories using an assemblage of technological platforms. This process of participatory infrastructuring furthered Grönvall et al's. (2016) claim that values are contextual and change based on the situation one is designing for. The value of supporting the community's data infrastructure literacy, while contextually relevant and asked for during the workshops contradicted the resource and time constrained in which many of the community leaders operated. Building the workforce training module in this context served as a case study in the participatory infrastructuring of civic data infrastructures. This process of participatory infrastructuring was grounded in principles of data and design justice (Bardzell, 2018; Costanza-Chock, 2020; D'Ignazio & Klein, 2020; Dombrowski et al., 2016) in that it prioritized local values and needs, built on technologies that were already working rather than building from scratch, and produced working prototypes that could be maintained by the community after researchers had left the field. This form of infrastructuring for social justice is a growing area of interest in the study of information infrastructures (Agid, 2018; Halkola et al., 2015; Korn et al., 2019; Prost et al., 2019; Vasilis Vlachokyriakos et al., 2018b) and one that I hope to contribute to going forward.

The human elements of this data infrastructure that I identified as well as the local conceptualizations of data equity are unique to the group of community leaders that I was working with and the sociopolitical context that shaped my engagement with them. The

findings are also a consequence of my own positionality within the data infrastructure, which was that of a student bound by the institutional requirements set by the university that employed me. My status as a student also limited the kinds of access I had within the university. There were many who had an influence on the early conceptualization of the dashboard or made decisions that impacted how the dashboard was funded, whom I did not have a chance to speak with directly. Some staff and faculty members I spoke with alluded to departmental politics as one of the reasons for the dashboard being structured the way it was, which I was not able to follow up on to confirm or deny.

While the general approach I followed during the infrastructural inversion of the dashboard, the organizing of the workshops and the participatory infrastructuring of the Westside's civic data can be generalized, the specifics would depend on the community one is designing for and with. Each community is defined by its own history, resources and relationships that will dictate how they engage and what they would value in their civic data infrastructure. Researchers and practitioners would be well served by customizing these methods to better suit the sociotechnical context in which they are engaging.

This is to say, success of such community projects depend to a large extent on the ability of the individuals leading the engagement. In addition to their individual abilities as researchers, facilitators and designers, their sociopolitical affiliations, agendas and identities will also be crucial to the success of these projects. Being familiar with the neighborhood and what she believed were their primary matters of concern, it was much easier for Dean Royster to launch projects and get the community's buy in, while I had to spend a considerable amount of time building my relationship with them. It was this

unfamiliarity, along with my desire to pursue a participatory design agenda that prompted me to rely on the expertise of the community leaders rather than my own. I imagine such an engagement would look a lot different when carried out by individuals within other institutions or organizations, where the relationship and power dynamics between them and the community are completely different.

In 2018 for instance, Neighborhood Nexus (NN) worked with the Westside Future Fund (WFF) to develop the Westside Community Data Dashboard (http://westsideprogress.org/). Not following the kind of infrastructural bricolage adopted by the CWK, NN developed this dashboard in a fraction of the time it took the CWK to develop its own dashboard. But the goals of the WFF's dashboard are very different from those of CWK. At an event organized by the WFF to promote the dashboard [17], presenters mentioned how the dashboard allowed corporate funders to see the impact their dollars were having on the Westside. Such a narrative will obviously be different from the one being forwarded by CWK and the one the community leaders are demanding. Identifying the values that animate such data infrastructures offer a starting point to consider what it takes to do justice to civic data that the communities can access and benefit from.

Data dashboards and civic infrastructures are being developed with promises being made about transparency, efficiency, equity and engagement in several cities across the US and the world. Future work would involve replicating my approach to infrastructural inversion and participatory infrastructuring across these cities and communities of practice to reveal other possible values and alignments between elements of their human

[17] https://www.westsidefuturefund.org/events/transform-westside-summit-friday-april-20-2018/

infrastructure. Insights from multiple such projects will allow me to develop a framework that researchers can use to ensure that future data dashboards and civic data infrastructures are grounded in a local understanding of values.

Aspects of this research could also be applied to gig workers, whose work is controlled by algorithmic systems like Uber, Mechanical Turk, InstaCart and the like, have been advocating for better employment policies and benefits. These systems have already been subject to scholarly criticism on the basis of how they treat their digital workers (M. L. Gray & Suri., 2019; Irani & Silberman, 2013). I would be interested in taking a participatory approach to unpacking the local values and infrastructural arrangements that underlie such algorithmic systems. Would a participatory infrastructuring approach that involved deploying working prototypes that prioritized local values and needs of these workers help produce more equitable employment policies and benefits? Would the artful integration of platforms or infrastructural bricolage help in the infrastructuring of socially just work environments? Other vulnerable communities like refugees are also being subject to algorithmic decision-making processes to determine the best cities that they could be relocated to in the United States (Trapp et al., 2018). How are values like justice, fairness, accountability and trust being considered in the design and implementation of such algorithmic and data infrastructures? Such questions would be relevant to researchers working in subfields of HCI like ICTD (Information and Communication Technologies for Development), CSCW (Computer Supported Collaborative Work and Social Media), PD, Community Informatics among others.

The United Nations has adopted Big Data as a key strategy in attaining its sustainable development goals, like peace, justice and strong institutions. Other countries have also recognized the value of data and taken it up as a key ingredient in their visions for socioeconomic development. Researchers concerned with using Information and Communication Technologies for Development (ICTD) are keen on not replicating western theories and models when designing data technologies for developing countries. They have developed postcolonial and feminist perspectives that are based in the lived experience of the communities they are designing for within these developing countries. My approach of developing a localized set of justice principles that were grounded in the experience of the community leaders I was designing for and with might be one avenue through which researchers and practitioners in the space of ICTD might also work with their communities.

Beyond the direct implications of this work for researchers and practitioners concerned with algorithmic and data infrastructures, elements of this work might be more broadly relevant to the study of Information Infrastructures in general. The conceptualization of data dashboards as windows into a community's underlying data infrastructure allows us to consider other such data assemblages that might serve as windows into their community's infrastructural arrangements. Based on my findings in this dissertation, I would argue that even simplistic data visualizations are mired in a series of sociotechnical interpretations which, if unpacked further can lead us to unpacking the layers of its underlying data infrastructure. Ultimately, I hope this dissertation helps researchers and practitioners move beyond the mere publicizing of data

as a strategy for empowerment, but instead think about realigning the human elements of the underlying data infrastructure in order to empower communities.

APPENDIX A. MEMORANDUM OF UNDERSTANDING

November 30, 2017

Firaz Peer
401, 16th St NW
Unit 1361
Atlanta, GA 30363

Dear Firaz,

I am pleased to partner with you to conduct focus group sessions for the Communities Who Know, Inc. Data Dashboard. For CWK, Inc., the focus group sessions are intended to support the upgrading process for the dashboard in creating a transportable prototype. For you, the sessions constitute a field project for your dissertation in the Ivan Allen College School of Literature, Media, and Communication. This mutually beneficial engagement is designed to address both of our needs in ways that are appropriate for both CWK goals and your educational goals.

Expectations

In keeping with our conversations about the focus group sessions:

- You will conduct 4 workshops on 1) data literacy and 2) the usability of the dashboard between December 2017 and March 2018, and conduct exit interviews between April and June 2018.
- You will keep me informed about operations and findings during this process so that the ongoing upgrading process for the data dashboard can be well-aligned with your discoveries, and you will be expected to participate, as appropriate, in CWK Dashboard team meetings.
- You will submit a final report on the focus group sessions and interviews to me by August 2018.
- You will also credit Communities Who Know, Inc. in your dissertation, and CWK, Inc. will cite your dissertation as a resource for the development of the CWK Dashboard.
- The CWK Project will provide a stipend for the community organizations that are participating in the focus group sessions. Operationally, what this means is:
    - The participants that you identify need to be representatives from community organizations in Westside Atlanta.

781 Marietta Street, Atlanta, GA 30318 PHONE 404.894.1727, FAX 404.894.8573, WEB iac.gatech.edu
A Unit of the University System of Georgia *An Equal Education and Employment Opportunity Institution*

  - Each organization will need to register through the Georgia Tech vendor system to receive payment upon completion of the project. Please check with Joanna Jeskova for specific instructions to them.
  - Each organization will receive that stipend at the completion of the 4 workshops and the interview.

- The CWK Project will provide light refreshments for each focus group session.
- After this project is completed, there will be a possibility of additional focus group sessions with Georgia Tech students and with high school students from Westside communities. The logistics of this arrangement will be worked through at the appropriate time but will include the possibility of a stipend for your services.

Professional Comportment and Confidentiality

The CWK Data Dashboard is my own Intellectual Property, the culmination of my years of research and action, my six-year partnerships with Westside communities of Atlanta (the Westside Communities Alliance, and the high quality work that the project team that I built conducted over the six-year period.

The conducting of the focus group sessions for data literacy and the usability of the CWK Data Dashboard will be the focus of your work in this collaboration. CWK will include you in team discussion for the dashboard as appropriate. Essentially, your focus group project will be contributing toward my work to establish a model for research in action. While the results of the focus groups will be available to you for the needs of your dissertation project as a field application for your desire to examine and analyze the "change in civic attitudes and behaviors of participants in data literacy workshops," you will not have direct access otherwise to the design and development of the CWK Dashboard or the right to present, publish, or share any elements of the dashboard in academic, professional, or other settings without my written permission. This contract confirms only my permission for you to feature the data from the workshops and interviews within your dissertation—with appropriate credit to CWK, Inc. Any additional uses that are not directly related to your dissertation requires my written permission. In other words, the expectation is that you will retain confidentially regarding the design and development of the CWK Dashboard, except as related to the data generated on data literacy by means of the collaboration on focus groups and the interviews specifically related to this project.

This agreement to confidentiality begins with your agreement to conduct the focus group sessions and the interviews and includes all future uses of CWK Data Dashboard information beyond your own dissertation project. By signing this contract, you agree to fulfill the expectations listed above and to maintain confidentiality and discretion about your contribution to my research project, the CWK Data Dashboard.

I look forward to working with you.

Sincerely,

Jacqueline J. Royster
Dean and
Executive Director
Communities Who Know, Inc.

I accept the terms of this agreement, agree to adhere to the expectations for the collaboration, and agree to maintain the specified professional comportment throughout the collaboration and for three years after the completion of the upgrade for the CWK Dashboard.

_____ _____

Signature: Firaz Peer Date

APPENDIX B. WORKSHOP FACILITATION WORKSHEETS

Build a Data Sculpture

Find a Story

Data is most useful when you can use it to tell a story about something. Using the data on this handout, try to find a story that you can tell by building a mini sculpture.

- Does one piece of data jump out at you?
- If you take a step back is there a pattern in the data?
- Do you see a story when comparing one part to another?

Make a Sculpture

Data visualization is very popular right now, but sometimes is hard to digest. Making a "sculpture" is a fun way to start playing with how to present your data story to other people.

- What symbols can you build to represent your data?
- How can you attract attention with this stuff?
- Can you tell simple and complex stories?

data BASIC.io

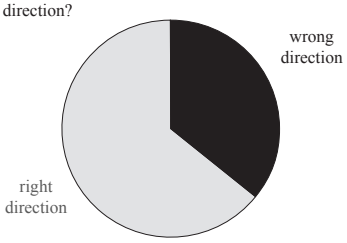
How Does Somerville Feel?

Happiness

Somerville asked its residents "How happy are you right now?"

The Future

Overall, would you say Somerville is moving in the right direction, or the wrong direction?



Sources:
- Somerville Wellbeing Survey

Build a Data Sculpture

Find a Story

Data is most useful when you can use it to tell a story about something. Using the data on this handout, try to find a story that you can tell by building a mini sculpture.

- Does one piece of data jump out at you?
- If you take a step back is there a pattern in the data?
- Do you see a story when comparing one part to another?

Make a Sculpture

Data visualization is very popular right now, but sometimes is hard to digest. Making a "sculpture" is a fun way to start playing with how to present your data story to other people.

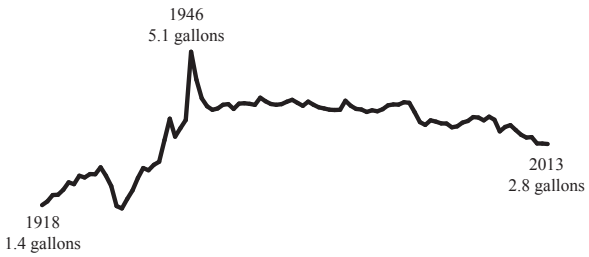
- What symbols can you build to represent your data?
- How can you attract attention with this stuff?
- Can you tell simple and complex stories?

data BASIC.io

Ice Cream In the US

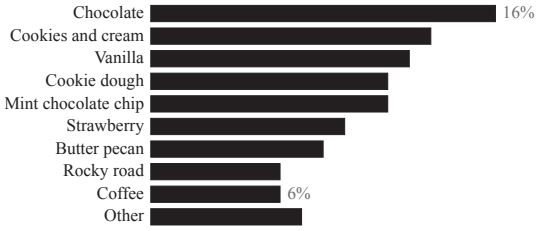
How Much Do We Eat?

This chart shows the gallons of ice cream an average person would eat in a year.



Our Favorite Flavors

This survey asked 1000 people what their favorite flavor of iceam was in 2014



Sources: USDA, Vission Critical

Build a Data Sculpture

Find a Story

Data is most useful when you can use it to tell a story about something. Using the data on this handout, try to find a story that you can tell by building a mini sculpture.

- Does one piece of data jump out at you?
- If you take a step back is there a pattern in the data?
- Do you see a story when comparing one part to another?

Make a Sculpture

Data visualization is very popular right now, but sometimes is hard to digest. Making a "sculpture" is a fun way to start playing with how to present your data story to other people.

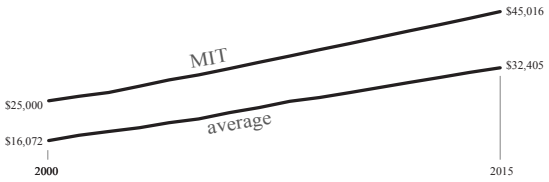
- What symbols can you build to represent your data?
- How can you attract attention with this stuff?
- Can you tell simple and complex stories?

data BASIC.io

How Much Does School Cost?

The Cost of an MIT Education

The total cost of tuition and fees at MIT and an "average" American undergraduate institution.



American Perceptions

Higher education institutions should reduce tuition and fees

State governments should provide more assistance

The federal government should provide more assistance

Companies should provide more assistance to employees

Sources:
- *America's Call for Higher Education Reform* Lumina Foundation and Gallup (2013)
- MIT Office of the Provost
- The College Board

word COUNTER

Sketch a Data Story

Use WordCounter to analyze a musician's lyrics and sketch a story with their words

What is WordCounter?

WordCounter analyzes text and tells you the most common words and phrases. It helps you quantitatively summarize your text so you can investigate how words are used. Sketching a data story helps participants build their data literacy by finding and drawing a story about what words show up in famous musicians' lyrics.

Learning Goals

- Increased ability to analyze and reason about text data.
- Understanding that one can find many stories in one dataset.
- Ability to quickly communicate a data-driven story to others.

Run the Activity

Solving a Problem

Text is data too! Finding patterns in large sets of text can be hard to do by hand, but computers can help. WordCounter helps you start to analyze text quantitatively but showing you the most commonly used words and phrases.

Share Inspirational Examples

Visualizing music lyrics is a popular thing to do. Show examples like Tahir Hemphill's "Rap Research Lab" (*http://rapresearchlab.com*) or Javier Arce's "Spotimap" (*http://javier.is/spotimap*). Ask participants what stories they see in these examples.

Introduce the Tool

Open up WordCounter (*https://databasic.io/wordcounter*) and choose "Elvis Presley" from the sample data drop-down menu. Press "Count!" to show an analysis of all of Elvis' lyrics. Show your group how the word cloud of his lyrics shows that he used the word "love" a lot, but only by looking at the bigrams and trigrams can you start to understand how and why. The fact that Elvis said "love" so times but only said "attorney" once could be a fun and revealing story to sketch. Mention that you can download CSVs of the results for further analysis.

Total time

30 to 45 Minutes

Audience

3 - 100 people. Ages 12+.

Designed for grades 6 - 12, Higher Ed classrooms, News Organizations, Non-profits, and Community Workshops. No prior experience with data is required.

Space

- A projector and computer.
- Ability to break out into small groups of 3 clustered around a computer.
- Large tables or floor, or tape to stick paper to walls so participants can draw

Supplies

- Computers
  1 for every 3 participants
- Large pieces of paper
  roughly 2 feet x 3 feet
- Thick crayons or markers

Brainstorm Questions

Use WTFcsv to identify some questions to ask your spreadsheet

What is WTFcsv?

WTFcsv gives you a quick overview of what is in your .csv file. It helps you quickly move to asking your spreadsheet some questions that can help you find stories to tell. This hands-on activity helps participants build their data literacy while looking at some fun datasets and brainstorming questions you can ask about them.

Learning Goals

- Understanding that data has many types, including numbers, text, and dates.
- Increased ability to form hypotheses about and find stories from metadata.
- Understanding that one can find many stories in one dataset.
- Understanding that datasets can be combined to answer more complicated questions.

Total time

30 to 45 Minutes

Audience

3 - 100 people. Ages 12+.
Designed for grades 6 - 12, Higher Ed classrooms, News Organizations, Non-profits, and Community Workshops. No prior experience with data is required.

Space

- A projector and computer.
- Ability to break out into small groups of 3 clustered around a computer.
- Large tables or floor, or tape to stick paper to walls so participants can draw

Supplies

- Computers
  1 for every 3 participants
- Copies of the WTFcsv question handout
  the third page of this guide
- Pens

Run the Activity

Solving a Problem

Open up a spreadsheet in Excel or something and poll the room for how many participants spend time working with data like this. Introduce the CSV file as simply one standard way to save spreadsheet data. Finding stories in a CSV can be hard, especially when there are lots of columns and rows. It's helpful to start by thinking about what questions you can even ask your data. WTFcsv helps you quickly analyze a CSV to understand what is in it so you can start asking it some questions.

Share Inspirational Examples

Show an open data portal from a place near the workshop location (for instance, *http://data.gov* in the US). Talk about the wide variety of datasets that are available online in portals like the one you are showing. Download our UFO sample data (*https://databasic.io/wtfcsv/ufo.csv*) and open it in Excel. Show that each column holds different types of information about each row.

Communities Who Know Data Dashboard
Understanding Data Workshop III
March 11, 2017
Information Hunt

The Communities Who Know Data Dashboard is a resource that has been developed to serve the data needs of the Westside communities and the rest of Atlanta. We hope this data will be helpful to those who are trying to use data to tell stories and make evidence based decisions.

Included here are some questions about data that can be found within the different modules of the CWK Data Dashboard. Please answer these questions by using the dashboard (available at http://communitieswhoknow.com/Home/) as a reference.

1. What is the percentage of the population in Atlanta that is age 5 to 17 years?
   _____

2. Compare the percent of individuals with only a bachelor's degree in NPU L with Atlanta.
   NPU L: _____
   Atlanta: _____

3. Atlanta Public School school with the highest 7-12 dropout rate and one with the lowest?
   Highest: _____
   Lowest: _____

4. Which schools have an "SAT High" score above 1400?
   _____
   _____

5. When was Herndon Homes was built:
   _____

6. When was Herndon Homes was demolished:
   _____

7. What is the percentage of the population in NPU K that is 65 years and over?
   _____

8. What are the annual expenditures of these schools for Instruction?
   Lin Elementary School: _____
   Bethune Elementary School: _____

9. When did construction begin on the MARTA rail system?
   _____

10. What is the percentage of remedial students in these schools?
    Grady High School: _____
    Booker T. Washington High School: _____
    KIPP Collegiate _____

11. In which year did AUC students march to protest segregation
    _____ and

    APS student march against desegregation at APS was in:
    _____

12. Compare the percent of the population that was white in 1940 and 2010 in NPU L?
    1940: _____
    2010: _____

Communities Who Know Data Dashboard

Understanding Data Workshop III
June 09, 2018

Information Hunt

The Communities Who Know Data Dashboard is a resource that has been developed to serve the data needs of the Westside communities and the rest of Atlanta. We hope this data will be helpful to those who are trying to use data to tell stories and make evidence based decisions.

Included here are some questions about data that can be found within the different modules of the CWK Data Dashboard. Please answer these questions by using the dashboard (available at http://communitieswhoknow.com/Home/) as a reference.

1. What is the percentage of the population in Atlanta that is age 5 to 17 years?
   _____

2. Compare the percent of individuals with only a bachelor's degree in NPU L with Atlanta.
   NPU L: _____
   Atlanta: _____

3. Atlanta Public School school with the highest 7-12 dropout rate and one with the lowest?
   Highest: _____
   Lowest: _____

4. Which schools have an "SAT High" score above 1400?
   _____
   _____

5. When was Herndon Homes was built:
   _____

6. When was Herndon Homes was demolished:
   _____

7. What is the percentage of the population in NPU K that is 65 years and over?
   _____

8. Which NPU had the most number of reported fire incidents?
   _____

9. What are the annual expenditures of these schools for Instruction?
   Lin Elementary School: _____
   Bethune Elementary School: _____

10. When did construction begin on the MARTA rail system?
    _____

11. What is the percentage of remedial students in these schools?
    Grady High School: _____
    Booker T. Washington High School: _____
    KIPP Collegiate _____

12. Approximately, how many crimes were reported close to Lindsay Street Baptist Church: _____

13. In which year did AUC students march to protest segregation
    _____ and

    APS student march against desegregation at APS was in:
    _____

14. Compare the percent of the population that was white in 1940 and 2010 in NPU L?
    1940: _____
    2010: _____

15. The total number of crimes and code violations reported between 2008 and 2015 for NPU L were
    Crimes: _____
    Code Violations: _____

Build Data Sculptures

The idea of playing with data is new to most people. This activity lets people quickly build sculptures that tell a simple data story with craft materials. The playful approach to the data helps engage the participants in thinking about how stories can be found and presented quickly and helps people feel more freedom and flexibility about data presentations.

How you do it:

1. Gather a collection of cheap craft materials – plastic bottle tops, wire, small fuzzy balls, markers, colored paper. Include ways to attach these together, like tape and glue.
2. Introduce the group to two related "normal" charts of data. One can be a single fact, and the other a medium-sized set of information.
3. Ask the group to pair up, preferably with someone they don't know.
4. Show participants a large central table full of the materials you have gathered. Give them 6 minutes to quickly build a physical representation of the data you presented earlier.
5. Stop everyone when the time is up.
6. Give each group 1 minute to share what they made.

Make Data Storybooks

Storytelling is an art form, and we don't get to practice it very much. This activity lets participants practice putting a data story together into a narrative, like a storyteller would. It lets people sketch their story and play with different ways to tell it in a fun storybook form, creating a narrative that can tell their stories in a convincing way.

How you do it:

1. Introduce a small set of data, and one story you have found in it.
2. Break the participants into pairs or threes, giving each a piece of paper and some pens. Have each group fold their piece of paper in half. Use big pieces of paper if you have them available.
3. Give the groups 10 minutes to sketch out a story on their paper. Encourage them to use drawings, words, shapes, and color creatively.
4. Bring everyone back together to share the stories they made.

Data Presentation Activities to Empower People

Here are short descriptions of four activities you can run that empower people to find and present a data story. Each activity includes a short description of why you would run it, and a step-by-step explanation.

Remix a Visualization Using Different Presentation Techniques

The goal of this activity is to practice the various techniques for presenting data. This gives participants a "toolbelt" of techniques they can use to tell a data story, helping them feel more confident that they can present data creatively.

How you do it:

1. Introduce an existing visualization and explain its audience, goals, and content.
2. Break the participants into groups of 3 or 4 people each. Assign each group a specific technique (personal story, data sculpture, map or creative map, chart or creative chart, data game). Give each group a copy of the visualization you introduced.
3. Let the groups work for 10 minutes to brainstorm and sketch one example that uses the same data in the visualization, but presents it using the technique you assigned to them.
4. Bring everyone back together and go around the room letting each group share their favorite idea.

Make Some Word Webs

Abstract ideas are hard to picture, and even harder to draw. A word web is a tool for exploring abstract ideas. This activity gives participants a way to turn abstract ideas into concrete images, allowing them to move from numbers to pictures to engage new audiences.

How you do it:

1. Spread out large pieces of paper, each with an abstract concept written in the middle.
2. Break the participants into group of 5 or 6. Each person should have a pen. Each group should have one of the pieces of paper you just showed.
3. Tell the participants they should start by drawing a line from the central word and writing another word that they associate with that one. Keep adding words connected to the first word or to the ones that other people add.
4. Give the groups 6 minutes to brainstorm and write words. You can force them to do this in silence, so they are focused on the words written down and nothing else.
5. Bring everyone back together, hang the sheets of paper on the wall, spend a few minutes letting everyone walk around looking at them to find words that inspire images.

Finding a "Factoid" Story

Sometimes in large sets of data you find the most interesting thing is the story of one particular piece of information. This could be an "outlier" (a data point not like the others), or it could be the data point that is most common. A detail about one particular piece of your data can fascinate and surprise people. It can also give them an easier way to start thinking about the whole set of data.

One factoid is that

_____

_____

_____

_____

This stands out from the rest of the data because

_____

_____

_____

_____

We want to tell this story because

_____

_____

_____

_____

Finding a "Interaction" Story

When two aspects of your data seem related, you can tell a story about how they interact. The fancy name for this is "correlation". If one measure goes up, the other goes up too. If one goes down, the other goes down. If other cases, they might interact as opposites (when one goes up, the other goes down). You need to be careful not to guess about reasons for the interaction, but noticing the relationship itself can be a good story that connects things people otherwise don't think about together.

The two pieces of the data that interact are

_____ _____

_____ and _____

_____ _____

The interaction is

_____

_____

_____

_____

We want to tell this story because

_____

_____

_____

_____

Finding a "Comparison" Story

Comparing between sections of your data can a good way to find a story to tell. Often one part of your data tells one story, but another part tells a totally different story. Or maybe there is a smaller portion of your data that serves as an example of an overall pattern.

The data to compare are

_____ _____

_____ and _____

_____ _____

Comparing these things shows that

_____

_____

_____

_____

We want to tell this story because

_____

_____

_____

_____

Finding a "Change" Story

People like to think about how things change over time. We experience and think about the world based on how we interact with it over time. Telling a story a story about change over time appeals to people's interest in understanding what causes change, and they can often remember seeing the differences.

The data show a change in

_____

_____

_____

_____

The data changed from

_____ _____

_____ **to** _____

_____ _____

We want to tell this story because

_____

_____

_____

_____

Finding a "Personal" Story

Some stories are interesting because they connect to your real life. Personalizing the story creates a connection to the real world meaning of the data and can be a powerful type of story for small audiences. Stories about someone's personal experiences can make the data seem more real.

The data say

_____

_____

_____

_____

This connects to real people because

_____

_____

_____

_____

We want to tell this story because

_____

_____

_____

_____

REFERENCES

Agar, M. H. (1996). *The Professional Stranger: An Informal Introduction to Ethnography*. Emerald Group Publishing Limited.

Agid, S. (2016). "□It's your project, but it's not necessarily your work□": Infrastructuring, situatedness, and designing relational practice. *Participatory Design Conference*. https://doi.org/10.1145/2940299.2940317

Agid, S. (2018). 'Dismantle, change, build': Designing abolition at the intersections of local, large-scale, and imagined infrastructures. *Design Studies*, *59*, 95–116. https://doi.org/10.1016/j.destud.2018.05.006

Ali, M., & Bailur, S. (2007). the Challenge of " Sustainability " in Ict4D – Is Bricolage the Answer ? *Proceedings of the 9th International Conference on Social Implications of Computers in Developing Coutries*.

Allen, M. (2017). The SAGE Encyclopedia of Communication Research Methods. In *The sage encyclopedia of communication research methods* (Vols. 1–4). SAGE Publications, Inc. https://doi.org/10.4135/9781483381411

Andrejevic, M. (2014). The big data divide. *International Journal of Communication*, *8*, 1673–1689. http://ijoc.org/index.php/ijoc/article/view/2161

Aoki, P. M., Honicky, R. J., Mainwaring, A., Myers, C., Paulos, E., Subramanian, S., & Woodruff, A. (2009). A Vehicle for Research: Using Street Sweepers to Explore the Landscape of Environmental Community Action. *ACM Transactions on Computer-Human Interaction*. https://doi.org/10.1145/1518701.1518762

Asad, M., Dantec, C. A. Le, Nielsen, B., & Diedrick, K. (2017). Creating a Sociotechnical API: Designing City-Scale Community Engagement. *In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2295–2306.

Aulakh, G., Alawadhi, N., & Agarwal, S. (2016). *Google in talks with government to partner for Aadhaar, UPI: Caesar Sengupta, Vice President, Next billion users at Google*. The Economic Times. https://economictimes.indiatimes.com/opinion/interviews/google-in-talks-with-government-to-partner-for-aadhaar-upi-caesar-sengupta-vice-president-next-billion-users-at-google/articleshow/54556320.cms

Baack, S. (2015). Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism. *Big Data & Society*, *2*(2).

Balestrini, M., Rogers, Y., Hassan, C., Creus, J., King, M., & Marshall, P. (2017). A City in common: A framework to orchestrate large-scale citizen engagement around urban issues. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3025453.3025915

Bannon, L., Bardzell, J., & Bødker, S. (2018). Introduction: Reimagining participatory design-Emerging voices. In *ACM Transactions on Computer-Human Interaction*. https://doi.org/10.1145/3177794

Bannon, L., & Ehn, P. (2013). Design Design Matters in Participatory Design. In J. Simonsen & T. Robertson (Eds.), *Routledge international handbook of participatory design* (pp. 37–63). Routledge.

Bannon, L. J. (1986). From Human Factors to Human Actors: The Role of Psychology and Human-Computer Interaction Studies in System Design. In J. M. Greenbaum & M. Kyng (Eds.), *Design at Work: Cooperative Design of Computer Systems* (pp. 25–44). Erlbaum. https://doi.org/10.1016/b978-0-08-051574-8.50024-8

Barab, S., & Squire, K. (2014). *Design-Based Research : Putting a Stake in the Ground*. *13*(1), 1–14.

Bardzell, S. (2010a). Feminist HCI : Taking Stock and Outlining an Agenda for Design. *Proceedings of the 28th International Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/1753326.1753521

Bardzell, S. (2010b). Feminist HCI. *Proceedings of the 28th International Conference on Human Factors in Computing Systems - CHI '10*. https://doi.org/10.1145/1753326.1753521

Bardzell, S. (2018). Utopias of participation: Feminism, design, and the futures. *ACM Transactions on Computer-Human Interaction*. https://doi.org/10.1145/3127359

Bartlett, J., & Tkacz., N. (2017). *Governance by dashboard: A policy paper*.

Baskett, L., LeRouge, C., & Tremblay, M. C. (2008). Using the dashboard technology

properly. *Health Progress*, *89.5*(16).

Bates, J., Lin, Y.-W., & Goodale, P. (2016). Data journeys: Capturing the socio-material constitution of data objects and flows. *Big Data & Society*, *3*(2), 205395171665450. https://doi.org/10.1177/2053951716654502

Benjamin, S., Bhuvaneswari, R., & P. Rajan, M. (2007). *Bhoomi: 'E-Governance',Or, An Anti-Politics Machine Necessary to Globalize Bangalore?*

Bhargava, R. (n.d.-a). *Finding a Story in Data*. https://datatherapy.org/activities/activity-finding-a-story-in-data/. https://datatherapy.org/activities/activity-finding-a-story-in-data/

Bhargava, R. (n.d.-b). *Remix a Visualization*. https://databasic.io/en/culture/remix. https://databasic.io/en/culture/remix

Bhargava, R. (2017). *You Don't Need a Data Scientist, You Need a Data Culture*. https://datatherapy.org/2017/12/06/building-a-data-culture/

Bhargava, R., Kadouaki, R., Bhargava, E., Castro, G., & D'Ignazio, C. (2016). Data Murals: Using the Arts to Build Data Literacy. *The Journal of Community Informatics*, *12*(3), 197–216.

Biehl, J. T., Czerwinski, M., Smith, G., & Robertson, G. G. (2007). FASTDash: a visual dashboard for fostering awareness in software teams. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1313–1322. https://doi.org/10.1145/1240624.1240823

Bietz, M. J., Baumer, E. P. S., & Lee, C. P. (2010). Synergizing in cyberinfrastructure development. *Computer Supported Cooperative Work*. https://doi.org/10.1007/s10606-010-9114-y

Bjerknes, G., & Bratteteig, T. (1987). Florence in Wonderland: System development with nurses. In G. Bjerknes, P. Ehn, & M. Kyng (Eds.), *Computers and Democracy: A Scandinavian Challenge* (pp. 279–295). Aldershot: Avebury.

Björgvinsson, E., Ehn, P., & Hillgren, P.-A. (2012). Design Things and Design Thinking: Contemporary Participatory Design Challenges. In *Design Issues* (Vol. 28, pp. 101–116). https://doi.org/10.1162/DESI_a_00165

Blomberg, J., Burrell, M., & Guest, G. (2003). An ethnographic approach to design. In *The HumanComputer Interaction Handbook* (pp. 964–986). https://doi.org/10.1201/b11963-52

Blomberg, J., Giacomi, J., Mosher, A., & Swenton-Wall, P. (1993). Ethnographic Field Methods and Their Relation to Design. *Participatory Design: Principles and Practices*, 123–155. https://doi.org/10.1007/BF01305846

Blomberg, J., & Karasti, H. (2013). Ethnography: Positioning ethnography within Participatory Design. In J. Simonsen & T. Robertson (Eds.), *Routledge international handbook of participatory design* (pp. 86–116). Routledge.

Bloom, M. (2016). *Atlanta school district to consider closing more schools*. The Atlanta Journal Constitution. https://www.ajc.com/news/local-education/atlanta-school-district-consider-closing-more-schools/4DjXfq13X6BADf1WZrKzQM/?fbclid=IwAR1DDsWjOukRzWoidw2UGSsonn5u1UKZPJ2yBLB446_SNIGai4BAJpjshgo

Blythin, S., Rouncefield, M., & Hughes, J. A. (1997). Never mind the ethno' stuff, what does all this mean and what do we do now: ethnography in the commercial world. *Interactions*. https://doi.org/10.1145/255392.255400

Bødker, S, Ehn, P., Kammersgaard, J., Kyng, M., & Sundblad, Y. (1987). A UTOPIAN experience: On design of powerful computer-based tools for skilled graphic workers. In *Computers and Democracy* (p. 25).

Bødker, Susanne. (2006). When second wave HCI meets third wave challenges. *Proceedings of the 4th Nordic Conference on Human-Computer Interaction: Changing Roles*, 1–8. https://doi.org/10.1145/1182475.1182476

Bødker, Susanne, Dindler, C., & Iversen, O. S. (2017). Tying Knots: Participatory Infrastructuring at Work. *Computer Supported Cooperative Work: CSCW: An International Journal*. https://doi.org/10.1007/s10606-017-9268-y

Bødker, Susanne, & Kyng, M. (2018). Participatory design that matters—Facing the big issues. *ACM Transactions on Computer-Human Interaction*. https://doi.org/10.1145/3152421

Boehner, K., & DiSalvo, C. (2016). Data, Design and Civics. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, 2970–2981.

https://doi.org/10.1145/2858036.2858326

Borning, A., & Muller, M. (2012). Next steps for Value Sensitive Design. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/2207676.2208560

Botchwey, N., Guhathakurta, S., & G. Zhang, G. (2014). *Atlanta's Neighborhood Quality of Life and Health Project. Center for Geographic Information systems*. http://www.cgis.gatech.edu/NQOLH/

Bowker, G. C. (1994). *Science on the run : information management and industrial geophysics at Schlumberger, 1920-1940*. MIT Press.

Bowker, G. C. (1996). The history of information infrastructures: The case of the International Classification of Diseases. In *Information Processing and Management*. https://doi.org/10.1016/0306-4573(95)00049-M

Bowker, G. C. (2013). Data Flakes: An Afterword to " Raw Data " Is an Oxymoron. In *Raw data is an oxymoron* (pp. 167–171).

Bowker, G. C. (2014). The infrastructural imagination. In A. Mongili & G. Pellegrino (Eds.), *Information Infrastructure(s): Boundaries, Ecologies, Multiplicity* (pp. xii–xiii). Cambridge Scholars Publishing.

Bowker, G. C., & Star, S. L. (1999). *Sorting Things Out: Classification and its Consequences*. MIT Press.

Boyd, D., & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society*, *15*(5), 662–679. https://doi.org/10.1080/1369118X.2012.678878

Britz, J., Hoffmann, A., Ponelis, S., Zimmer, M., & Lor, P. (2013). On considering the application of Amartya Sen's capability approach to an information-based rights framework. *Information Development*, *29*(2), 106–113. https://doi.org/10.1177/0266666912454025

Büscher, M., Gill, S., Mogensen, P., & Shapiro, D. (2001). Landscapes of practice: Bricolage as a method for situated design. *Computer Supported Cooperative Work*. https://doi.org/10.1023/A:1011293210539

Calhoun, C. (1992). *Habermas and the Public Sphere*. MIT Press. https://doi.org/10.1177/1742766507078417

Carroll, J. (2013). Human Computer Interaction - brief intro. In J. Lowgren, J. M. Carroll, M. Hassenzahl, & T. Erickson (Eds.), *The Encyclopedia of Human-Computer Interaction* (2nd ed.). https://doi.org/10.1177/0011392199047004007

Carroll, J., Beck, J., Dhanorkar, S., Binda, J., Gupta, S., & Zhu, H. (2018). Strengthening community data: towards pervasive participation. *Proceedings of the 19th Annual International Conference on Digital Government Research Governance in the Data Age - Dgo '18*, 1–9. https://doi.org/10.1145/3209281.3209312

Carroll, J., Hoffman, B., Han, K., & Rosson, M. B. (2015). Reviving community networks: hyperlocality and suprathresholding in Web 2.0 designs. *Personal and Ubiquitous Computing*, *19*(2), 477–491. https://doi.org/10.1007/s00779-014-0831-y

Carroll, J. M., & Rosson, M. B. (1996). Developing the Blacksburg Electronic Village. *Communications of the ACM*. https://doi.org/10.1145/240483.240498

Carroll, J., & Rosson, M. (2007). Participatory design in community informatics. *Design Studies*, *28*(3), 243–261. https://doi.org/10.1016/j.destud.2007.02.007

Carroll, J., & Rosson, M. (2013). Wild at Home: The Neighborhood as a Living Laboratory for HCI. *ACM Transactions on Computer-Human Interaction*, *20*(3), 1–28. https://doi.org/10.1145/2491500.2491504

Carroll, J., Shih, P. C., & Kropczynski, J. (2015). Community informatics as innovation in sociotechnical infrastructures. *The Journal of Community Informatics*.

Chamberlain, A., Crabtree, A., Rodden, T., Jones, M., & Rogers, Y. (2012). Research in the wild: Understanding "in the wild" approaches to design and development. *Proceedings of the Designing Interactive Systems Conference, DIS '12*. https://doi.org/10.1145/2317956.2318078

Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis (Introducing Qualitative Methods series)*. Sage. https://doi.org/10.1080/17482620600881144

Ciborra, C. (2007). The Labyrinths of Information: Challenging the Wisdom of Systems.

In *The Labyrinths of Information: Challenging the Wisdom of Systems*. https://doi.org/10.1093/acprof:oso/9780199275267.001.0001

Ciborra, C. U. (1992). From thinking to tinkering: The grassroots of strategic information systems. *Information Society*. https://doi.org/10.1080/01972243.1992.9960124

Clarke, A. E. (2003). Situational Analyses: Grounded Theory Mapping After the Postmodern Turn. *Symbolic Interaction*, *26*(4), 553–576. https://doi.org/10.1525/si.2003.26.4.553

Clarke, C. (2009). *Storytelling for grantseekers: A guide to creative nonprofit fundraising*. John Wiley & Sons.

Cohen, R. L., & Greenberg, J. (1982). The Justice Concept in Social Psychology. In *Equity and Justice in Social Behavior*. https://doi.org/10.1016/b978-0-12-299580-4.50007-0

Collins, P. (2000). Black feminist thought: Knowledge, consciousness, and the politics of empowerment (perspectives on gender). In *Routledge*.

Corbin, J., & Strauss, A. (2008). *Basics of Qualitative Research (3rd ed.): Techniques and Procedures for Developing Grounded Theory*. SAGE Publications, Inc. https://doi.org/10.4135/9781452230153

Costanza-Chock, S. (2018). Design Justice: towards an intersectional feminist framework for design theory and practice. *DRS2018: Catalyst*. https://doi.org/10.21606/drs.2018.679

Costanza-Chock, S. (2020). *Design justice: Community-led practices to build the worlds we need*. MIT Press. https://doi.org/10.21606/drs.2018.679

Crabtree, A. (1998). Ethnography in Participatory Design. *Proceedings of the 1998 Participatory Design Conference*.

Crabtree, A., Rodden, T., Tolmie, P., & Button, G. (2009). Ethnography considered harmful. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/1518701.1518835

Crain, M., Poster, W., & Cherry, M. (2016). *Invisible Labor: Hidden Work in the Contemporary World*. University of California Press.

Crawford, K., Miltner, K., & Gray, M. L. (2014). Critiquing Big Data: Politics, Ethics, Epistemology. *International Journal of Communication*, *8*, 1663–1672. http://ijoc.org/index.php/ijoc/article/view/2167/1164

Crenshaw, K. (2018). Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory, and antiracist politics [1989]. In *Feminist Legal Theory: Readings in Law and Gender*. https://doi.org/10.4324/9780429500480

Crivellaro, C., Taylor, A., Vlachokyriakos, V., Comber, R., Nissen, B., & Wright, P. (2016). Re-Making Places: HCI , "Community Building" and Change. *CHI'16, May 07-12, 2016, San Jose, CA, USA*. https://doi.org/10.1145/2858036.2858332

D'Ignazio, C., & Bhargava, R. (n.d.-a). *Data Culture*. https://databasic.io/en/culture/. https://databasic.io/en/culture/

D'Ignazio, C., & Bhargava, R. (n.d.-b). *Data Sculpture Activity*. https://databasic.io/en/culture/build-a-sculpture. https://databasic.io/en/culture/build-a-sculpture

D'Ignazio, C., & Bhargava, R. (2016). DataBasic: Design Principles, Tools and Activities for Data Literacy Learners. *The Journal of Community Informatics*, *12*(3).

D'Ignazio, C., & Klein, L. F. (2016). Feminist Data Visualization. *Workshop on Visualization for the Digital Humanities (VIS4DH), Blati*.

D'Ignazio, C., & Klein, L. F. (2020). *Data Feminism*. MIT Press.

Dalton, J., & Thatcher, J. (2014). What does a critical data studies look like , and why do we care? Seven points for a critical approach to 'big data.' *Society and Space*, 1–12. http://societyandspace.com/material/commentaries/craig-dalton-and-jim-thatcher-what-does-a-critical-data-studies-look-like-and-why-do-we-care-seven-points-for-a-critical-approach-to-big-data/

De Certeau, M., & Mayol, P. (1998). The Practice of Everyday Life II: Living and Cooking. In *Living and cooking* (Vol 2). University of Minnesota Press.

Dencik, L., Hintz, A., & Cable, J. (2016). Towards data justice? The ambiguity of anti-surveillance resistance in political activism. *Big Data and Society*. https://doi.org/10.1177/2053951716679678

Dencik, L., Hintz, A., Redden, J., & Treré, E. (2019). Exploring Data Justice: Conceptions, Applications and Directions. *Information Communication and Society*. https://doi.org/10.1080/1369118X.2019.1606268

Dewey, J. (1927). *The Public and Its Problems* (1954th ed.). Swallow Press.

Dick, B. (2000). *Grounded theory: a thumbnail sketch*. https://doi.org/10.1109/DATE.2008.4484729

Dickinson, J., Díaz, M., Erete, S., & Riedl, D. L. (2018). Inclusion of Underserved Residents in City Technology Planning. *2018 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3170427.3188583

Dillahunt, T. R., & Veinot, T. C. (2018). Getting there: Barriers and facilitators to transportation access in underserved communities. *ACM Transactions on Computer-Human Interaction*. https://doi.org/10.1145/3233985

DiSalvo, C. (2009). Design and the Construction of Publics. *Design Issues*. https://doi.org/10.1162/desi.2009.25.1.48

DiSalvo, C., Clement, A., & Pipek, V. (2012). Participatory Design For, With, and By Communities. In J. Simonsen & T. Robertson (Eds.), *International Handbook of Participatory Design* (pp. 182–209). Routledge.

Disalvo, C., & Lukens, J. (2009). Towards a Critical Technological Fluency : The Confluence of Speculative Design and Community Technology Programs. *Proceedings of the Digital Arts and Culture Conference, 2009: After Media: Embodiment and Context*, 1–6.

Dombrowski, L., Harmon, E., & Fox, S. (2016). Social justice-oriented interaction design: Outlining key design strategies and commitments. *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS)*. https://doi.org/10.1145/2901790.2901861

Dörk, M., Feng, P., Collins, C., & Carpendale, S. (2013). Critical InfoVis: Exploring the

Politics of Visualization. *CHI '13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA '13*. https://doi.org/10.1145/2468356.2468739

Dourish, P. (2004). What we talk about when we talk about context. *Personal and Ubiquitous Computing*, *8*(1), 19–30. https://doi.org/10.1007/s00779-003-0253-8

Dourish, P. (2014). Reading and Interpreting Ethnography. In J. S. Olson & W. A. Kellogg (Eds.), *Ways of Knowing in HCI* (pp. 1–23). Springer. https://doi.org/https://doi-org.prx.library.gatech.edu/10.1007/978-1-4939-0378-8_1

Dourish, P. (2006). Implications for design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '06*, 541. https://doi.org/10.1145/1124772.1124855

Eck, J., Chainey, S., Cameron, J., & Wilson, R. (2005). *Mapping crime: Understanding hotspots*. National Institute of Justice.

Edwards, D., & Thomas, J. C. (2005). Developing a municipal performance-measurement system: Reflections on the Atlanta Dashboard. *Public Administration Review*, *65*(3), 369–376. https://doi.org/10.1111/j.1540-6210.2005.00461.x

Edwards, P., Bowker, G., Jackson, S., & Williams, R. (2009). Introduction: An Agenda for Infrastructure Studies. *Journal of the Association for Information Systems*. https://doi.org/10.17705/1jais.00200

Edwards, P. N. (2010). *A vast machine: Computer models, climate data, and the politics of global warming*. MIT Press.

Edwards, P. N., Jackson, S. J., Bowker, G. C., & Knobel, C. P. (2007). Understanding Infrastructure: Dynamics, Tensions, and Design: Report of a Workshop on History & Theory of Infrastructure, Lessons for New Scientific Cyberinfrastructures. In *Understanding Infrastructure Dynamics Tensions and Design Report of a Workshop on History Theory of Infrastructure Lessons for New Scientific Cyberinfrastructures*.

Edwards, P. N., Jackson, S. J., Chalmers, M. K., Bowker, G. C., Borgman L., C., Ribes, D., Burton, M., & Calvert, S. (2013). *Knowledge Infrastructures: Intellectual Frameworks and Research Challenges*.

Ehn, P. (2008). Participation in design things. *Conference on Participatory Design*, 92–

101. http://portal.acm.org/citation.cfm?id=1795248

Ehn, P. (2017). Learning in Participatory Design As I Found It (1970-2015). In B. Disalvo, J. Yip, E. Bonsignore, & C. Disalvo (Eds.), *Participatory Design for Learning: Perspectives from Practice and Research* (pp. 7–21). Routledge.

Ehn, P., & Sandberg., Å. (1983). *Local union influence on technology and work organization: Some results from the DEMOS project*. Arbetslivscentrum.

Erete, S., & Burrell, J. O. (2017). Empowered participation: Exploring how citizens use technology in local governance. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3025453.3025996

Erete, S., Israni, A., & Dillahunt, T. (2018). An intersectional approach to designing in the margins. *Interactions*. https://doi.org/10.1145/3194349

Erete, S., Ryou, E., Smith, G., Fassett, K. M., & Duda, S. (2016). Storytelling with Data: Examining the Use of Data by Non-Profit Organizations. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW '16*, 1271–1281. https://doi.org/10.1145/2818048.2820068

Erickson, I., & Sawyer, S. (2019). Infrastructuring as Bricolage: Thinking like a Contemporary Knowledge Worker. In M. M. Kornberger, G. Bowker, J. Elyachar, A. Mennicken, P. Miller, J. R. Nucho, & N. Pollock (Eds.), *Knowledge Infrastructures* (pp. 315–333). Springer.

Few, S. (2006a). Common Pitfalls in Dashboard Design. In *ProClarity Corporation*. https://www.perceptualedge.com/articles/Whitepapers/Common_Pitfalls.pdf

Few, S. (2006b). Information Dashboard Design. *The Effective Visual Communication of Data Sebastopol*, *January*, 223. https://doi.org/10.1017/S0021849904040334

Flanagan, M., Howe, D. C., & Nissenbaum, H. (2008). Embodying values in technology: Theory and practice. In *Information Technology and Moral Philosophy*. https://doi.org/10.1017/CBO9780511498725.017

Fox, S., Asad, M., Lo, K., Dimond, J. P., Dombrowski, L. S., & Bardzell, S. (2016). Exploring Social Justice, Design, and HCI. *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA*

'16. https://doi.org/10.1145/2851581.2856465

Frank, M., & Walker, J. (2016). Some Key Challenges for Data Literacy. *The Journal of Community Informatics*, *12*(3).

Frank, M., Walker, J., Attard, J., & Tygel, A. (2016). Data Literacy-What is it and how can we make it happen? *The Journal of Community Informatics*, *12*(3).

Fraser, N. (1990). Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy. *Social Text*. https://doi.org/10.2307/466240

Freire, P. (2000). *Pedagogy of the oppressed*. Bloomsbury Publishing.

Friedman, B., Kahn, P. H., & Borning, A. (2009). Value Sensitive Design and Information Systems. In *The Handbook of Information and Computer Ethics*. https://doi.org/10.1002/9780470281819.ch4

Glaser, B. G., & Strauss, A. L. (1967). *The Discovery of Grounded Theory*.

Goldsmith, S., & Crawford, S. (2014). *The responsive city: Engaging communities through data-smart governance*. John Wiley & Sons.

Gonzales, A. (2017). Technology maintenance: A new frame for studying poverty and marginalization. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3025453.3025475

González-Zapata, F., & Heeks, R. (2015). Understanding Multiple Roles of Intermediaries in Open Government Data. *13th International Conference on Social Implications of Computers in Developing Countries*. https://doi.org/10.1108/00220410510578005

Gordon, E., & Mihailidis, P. (2016). *Civic Media: Technology, Design, Practice*. MIT Press.

Gray, J., Gerlitz, C., & Bounegru, L. (2018). Data infrastructure literacy. *Big Data & Society*, *5*(2), 205395171878631. https://doi.org/10.1177/2053951718786316

Gray, M. L., & Suri., S. (2019). *Ghost Work: How to Stop Silicon Valley from Building a*

New Global Underclass (E. D. Books (ed.)).

Grönvall, E., Malmborg, L., & Messeter, J. (2016). Negotiation of values as driver in community-based PD. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/2940299.2940308

Grudin, J. (1993). Obstacles to participatory design in large product development organizations. In D. Schuler & A. Namioka (Eds.), *Participatory Design: Principles and Practices* (1st Editio). CRC Press. https://doi.org/10.1201/9780203744338

Gurstein, M. (2003). Effective use: A community informatics strategy beyond the digital divide. *First Monday*, *8*(12). https://doi.org/10.5210/fm.v8i12.1107

Gurstein, M. (2007). What is Community Energy & Why Does it Matter? In *Polimetrica*.

Gurstein, M. (2011). Open data: Empowering the empowered or effective data use for everyone? *First Monday*, *16*(2), 1–8. https://doi.org/10.1177/0170840601223003

Habermas, J. (1991). *The structural transformation of the public sphere: An inquiry into a category of bourgeois society*. MIT press.

Halkola, E., Iivari, N., & Kuure, L. (2015). Infrastructuring as social action. *2015 International Conference on Information Systems: Exploring the Information Frontier, ICIS 2015*.

Halse., J., Brandt, E., Clarke, B., & Binder, T. (2010). *Rehearsing the Future*. Danish Design School Press.

Hanseth, O., & Lyytinen, K. (2016). Design theory for dynamic complexity in information infrastructures: The case of building internet. In *Enacting Research Methods in Information Systems: Volume 3*. https://doi.org/10.1007/978-3-319-29272-4_4

Hanseth, O., & Monteiro, E. (1998). Understanding Information Infrastructure. In *Unpublished*.

Hara, K., Le, V., & Froehlich, J. (2013). Combining crowdsourcing and google street

view to identify street-level accessibility problems. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. https://doi.org/10.1145/2470654.2470744

Haraway, D. (1988). Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies*, *14*(3), 575. https://doi.org/10.2307/3178066

Harding, S. (Ed.). (2004). *The Feminist Standpoint Theory Reader: Intellectual and Political Controversies*. Routledge.

Harkavy, I. (2006). The role of universities in advancing citizenship and social justice in the 21st century. *Education, Citizenship and Social Justice*, *1*(1), 5–37. https://doi.org/10.1177/1746197906060711

Harper, R. H., & Hughes., J. A. (1993). What a F-ing System?: Send'em All to the Same Place and Then Expect Us to Stop'em Hitting; Making Technology Work in Air Traffic Control. Rank Xerox EuroPARC, 1991. In G. Button (Ed.), *Technology in Working Order: Studies of Work, Interaction, and Technology* (pp. 127–144). Routledge.

Heeks, R., & Renken, J. (2018). Data justice for development: What would it mean? *Information Development*. https://doi.org/10.1177/0266666916678282

Heeks, R., & Shekhar, S. (2019). Datafication, development and marginalised urban communities: an applied data justice framework. *Information Communication and Society*. https://doi.org/10.1080/1369118X.2019.1599039

Hughes, J. A., Randall, D., & Shapiro, D. (1992). From ethnographic record to system design - Some experiences from the field. *Computer Supported Cooperative Work*, *1*(3), 123–141. https://doi.org/10.1007/BF00752435

Hughes, J., King, V., Rodden, T., & Andersen, H. (1994). Moving Out from the Control Room: Ethnography in System Design. *Sociology The Journal Of The British Sociological Association*, *teractions*(4), 429–439. https://doi.org/10.1145/192844.193065

Hughes, T. P. (1984). Hughes, T. P. (1984). Networks of Power: Electrification in Western Society. In *Technology and Culture*. https://doi.org/10.1089/ten.tec.2010.0320

Huh, J., Nathan, L. P., Silberman, S., Blevis, E., Tomlinson, B., Sengers, P., & Busse, D. (2010). Examining appropriation, re-use, and maintenance for sustainability. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/1753846.1754173

Irani, L. C., & Silberman, M. S. (2013). Turkopticon: Interrupting worker invisibility in Amazon Mechanical Turk. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/2470654.2470742

Iversen, O., Halskov, K., & Leong, T. (2012). Values-led participatory design. *CoDesign*, *8*(September), 87–103.

Jackson, S. J., Edwards, P. N., Bowker, G. C., & Knobel, C. P. (2007). Understanding infrastructure: History, heuristics, and cyberinfrastructure policy. *First Monday*. https://doi.org/10.5210/fm.v12i6.1904

Jackson, S. J., Pompe, A., & Krieshok, G. (2012). Repair worlds: Maintenance, repair, and ICT for development in rural Namibia. *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*. https://doi.org/10.1145/2145204.2145224

Jacobson, B. R. (2017). *Leaving No One Behind in the Data Revolution: An analysis of four years of High-Level Political Forum reports*.

JafariNaimi (Parvin), N., Nathan, L. P., & Hargraves, I. (2015). Values as Hypotheses: Design, Inquiry, and the Service of Values. *Design Issues*, *31*(4).

Johnson, J. (2014). From open data to information justice. *Ethics and Information Technology*. https://doi.org/10.1007/s10676-014-9351-8

Johnson, J. A. (2016). *The value–and limits–of distributive justice in a justice-centered approach to information privacy*. In Western Political Science Association 2016 Annual Meeting.

Kahn, A., Nelson, R. E., Gaeddert, W. P., & Hearn, J. L. (1982). The Justice Process: Deciding Upon Equity or Equality. *Social Psychology Quarterly*. https://doi.org/10.2307/3033668

Kandel, S., Paepcke, A., Hellerstein, J. M., & Heer, J. (2012). Enterprise data analysis

and visualization: An interview study. *IEEE Transactions on Visualization and Computer Graphics*. https://doi.org/10.1109/TVCG.2012.219

Karasti, H. (2014). Infrastructuring in participatory design. *Proceedings of the 13th Participatory Design Conference on Research Papers - PDC '14*. https://doi.org/10.1145/2661435.2661450

Karasti, H., Baker, K. S., & Halkola, E. (2006). Enriching the notion of data curation in e-Science: Data managing and information infrastructuring in the Long Term Ecological Research (LTER) network. *Computer Supported Cooperative Work*, *15*(4), 321–358. https://doi.org/10.1007/s10606-006-9023-2

Karasti, H., & Blomberg, J. (2018). Studying Infrastructuring Ethnographically. *Computer Supported Cooperative Work: CSCW: An International Journal*, *27*(2), 233–265. https://doi.org/10.1007/s10606-017-9296-7

Karasti, H., & Syrjänen, A. L. (2004). Artful infrastructuring in two cases of community PD. *PDC 2004 - Proceedings of the Eight Participatory Design Conference 2004 - Artful Integration: Interweaving Media, Materials and Practices*. https://doi.org/10.1145/1011870.1011874

Kayser-Bril, N. (2016). Don't ask too much from data literacy. *The Journal of Community Informatics*, *12*(3).

Kensing, F., & Greenbaum, J. (2013). Heritage: having a say. In J. Simonsen & T. Robertson (Eds.), *Routledge international handbook of participatory design* (pp. 21–36). Routledge.

Kensing, F., Simonsen, J., & Bodker, K. (1998). MUST: A Method for Participatory Design. *Human-Computer Interaction*, *13*(2), 167–198. https://doi.org/10.1207/s15327051hci1302_3

Kensing, F., Simonsen, J., & Bødker, K. (1998). Participatory Design at a Radio Station. *Computer Supported Cooperative Work*. https://doi.org/10.1023/A:1008683004336

Kirkpatrick, R., & Vacarelu, F. (2018). A Decade of Leveraging Big Data for Sustainable Development. *UN Chronicle*. https://unchronicle.un.org/article/decade-leveraging-big-data-sustainable-development

Kitchin, R. (2014). The real-time city? Big data and smart urbanism. *GeoJournal*, *79*(1), 1–14. https://doi.org/10.1007/s10708-013-9516-8

Kitchin, R., & Lauriault, T. P. (2014). Towards critical data studies : Charting and unpacking data assemblages and their work. *Geoweb and Big Data*, 1–19. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2474112

Kitchin, R., Lauriault, T. P., & McArdle, G. (2015). Knowing and governing cities through urban indicators, city benchmarking and real-time dashboards. *Regional Studies, Regional Science*, *2*(1), 6–28. https://doi.org/10.1080/21681376.2014.983149

Kitchin, R., Maalsen, S., & McArdle, G. (2015). The Praxis and Politics of Building Urban Dashboards. *SSRN 2608988*.

Korn, M., Reißmann, W., Röhl, T., & Sittler, D. (2019). Infrastructuring Publics: A Research Perspective. In M. Korn, W. Reißmann, T. Röhl, & D. Sittler (Eds.), *Infrastructuring Publics* (pp. 11–48). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-20725-0_2

Korn, M., & Voida, A. (2015). Creating Friction: Infrastructuring Civic Engagement in Everyday Life. *Aarhus Series on Human Centered Computing*, *1*(1), 12. https://doi.org/10.7146/aahcc.v1i1.21198

Kruse, K. M. (2019). What does a traffic jam in Atlanta have to do with segregation? Quite a lot. *The New York Times*. https://www.nytimes.com/interactive/2019/08/14/magazine/traffic-atlanta-segregation.html

Kubitschko, S. (2015). Hackers media practices: Demonstrating and articulating expertise as interlocking arrangements. *Convergence*. https://doi.org/10.1177/1354856515579847

Kumar, R., & Best, M. L. (2006). Impact and sustainability of E-government services in developing countries: Lessons learned from Tamil Nadu, India. *Information Society*, *22*(1), 1–12. https://doi.org/10.1080/01972240500388149

Kurgan, L. (2013). Million Dollar Blocks. *Close Up at a Distance: Mapping, Technology, and Politics*.

Kyng, M., & Mathiassen, L. (1980). Systems Development and Trade Union Activities. *DAIMI Report Series*. https://doi.org/10.7146/dpb.v8i99.6515

Larkin, B. (2013). The politics and poetics of infrastructure. *Annual Review of Anthropology*, *42*, 327-343.

Latonero, M., & Gold, Z. (2016). *Data, Human Rights & Human Security*.

Latour, B. (2004). Why Has Critique Run out of Steam? From Matters of Fact to Matters of Concern. *Critical Inquiry*. https://doi.org/10.2307/1344358

Latour, B. (2005). From Realpolitik to Dingpolitik or How to Make Things Public. *Making Things Public-Atmospheres of Democracy*, 4–31.

Latour, B. (2008). A Cautious Prometheus? A Few Steps Toward a Philosophy of Design (with Special Attention to Peter Sloterdijk). *Design History Society*, 2. http://books.google.com/books?hl=en&lr=&id=3B5bbv0r-EkC&oi=fnd&pg=PA2&dq=A+Cautious+Prometheus?+A+Few+Steps+Toward+a+Philosophy+of+Design+(with+Special+Attention+to+Peter+Sloterdijk&ots=BENSsMyS03&sig=6iebJo1hrW2idy0DTtcUUscMi2

Le Dantec, C. A., Asad, M., Misra, A., & Watkins, K. E. (2015). Planning with Crowdsourced Data. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15*, 1717–1727. https://doi.org/10.1145/2675133.2675212

Le Dantec, C. A., & Edwards, W. K. (2008). Designs on dignity. *Proceeding of the Twenty-Sixth Annual CHI Conference on Human Factors in Computing Systems - CHI '08*, 627. https://doi.org/10.1145/1357054.1357155

Le Dantec, C. A., Poole, E. S., & Wyche, S. P. (2009). Values as lived experience: Evolving value sensitive design in support of value discovery. *Proceedings of the 27th International Conference on Human Factors in Computing Systems*, 1141–1150. https://doi.org/10.1145/1518701.1518875

Le Dantec, C., & DiSalvo, C. (2013). Infrastructuring and the formation of publics in participatory design. *Social Studies of Science*. https://doi.org/10.1177/0306312712471581

Le Dantec, C., & Fox, S. (2015). Strangers at the Gate: Gaining Access, Building Rapport, and Co-Constructing Community-Based Research. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 1348–1358. https://doi.org/10.1145/2675133.2675147

Lee, C. P., Dourish, P., & Mark, G. (2006). The Human Infrastructure of Cyberinfrastructure. *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work CSCW 06*. https://doi.org/10.1145/1180875.1180950

Letouzé, E., Bhargava, R., Deahl, E., Noonan, A., Sangokoya, D., & Shoup, N. (2015). Beyond data literacy: reinventing community engagement and empowerment in the age of data. In *Data-Pop Alliance White paper series* (p. 42 pages).

Levi-Strauss, C. (1966). *The Savage Mind*. University of Chicago Press.

Lippmann, W. (1993). *The Phantom Public*. Transaction Publishers.

Locke, K. (2000). *Grounded Theory in Management Research*. Sage. https://doi.org/10.4135/9780857024428

Lodha, S. K., & Verma, A. K. (2000). Spatio-temporal visualization of urban crimes on a GIS grid. *Proceedings of the ACM Workshop on Advances in Geographic Information Systems*.

Loukissas, Y. (n.d.). *Civic Data Guide: Thinking Critically About Digital Public Records*. https://serve-learn-sustain.gatech.edu/civic-data-guides-thinking-critically-about-digital-public-records

Loukissas, Y. (2016). Taking Big Data apart: local readings of composite media collections. *Information, Communication & Society*, *4462*(August), 1–14. https://doi.org/10.1080/1369118X.2016.1211722

Loukissas, Y. (2019). *All Data Are Local : Thinking Critically in a Data-Driven Society*. MIT Press. https://mitpress.mit.edu/books/all-data-are-local

Lukens, J. (2013). DIY Infrastructure and the Scope of Design Practice. *Design Issues*, *29*(3), 14–27. https://doi.org/10.1162/desi_a_00218

Manzini, E. (2015). *Design, When Everybody Designs: An Introduction to Design for Social Innovation*.

Marttila, S., & Botero, A. (2017). Infrastructuring for Cultural Commons. *Computer Supported Cooperative Work: CSCW: An International Journal*. https://doi.org/10.1007/s10606-017-9273-1

Maskell, T., Crivellaro, C., Anderson, R., Nappey, T., Araújo-Soares, V., & Montague, K. (2018). Spokespeople: Exploring Routes to Action through Citizen-Generated Data. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. https://doi.org/10.1145/3173574.3173979

Mattern, S. (2014). Interfacing Urban Intelligence. *Places Journal*, *April 2014*. https://doi.org/10.1017/CBO9781107415324.004

Mattern, S. (2015). Mission Control: A History of the Urban Dashboard. *Places*, *1*, 1–31. https://doi.org/10.1017/CBO9781107415324.004

Meng, A., & DiSalvo, C. (2018). Grassroots resource mobilization through counter-data action. *Big Data & Society*. https://doi.org/10.1177/2053951718796862

Mihailidis, P. (2018). Civic media literacies: re-Imagining engagement for civic intentionality*. *Learning, Media and Technology*. https://doi.org/10.1080/17439884.2018.1428623

Millen, D. R. (2000). Rapid ethnography: Time deepening strategies for HCI field research. *Proceedings of the Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, DIS*.

Monk, L., & Tutt, E. (n.d.). *Storytelling with Data*. https://www.neighborhoodindicators.org/data-tech/course-catalog/data-101-data-visualization-data-literacy-and-storytelling

Mörtberg, C., Bratteteig, T., Wagner, I., Stuedahl, D., & Morrison, A. (2010). Methods That Matter in Digital Design Research. In I. Wagner, T. Bratteteig, & D. Stuedahl (Eds.), *Exploring Digital Design: Multi-Disciplinary Design Practices*. Springer-Verlag. https://doi.org/10.1007/978-1-84996-223-0_4

Muller, M. (2014). Curiosity, Creativity, and Surprise as Analytic Tools: Grounded

Theory Method. In *Ways of Knowing in HCI* (pp. 25–48). Springer New York. https://doi.org/10.1007/978-1-4939-0378-8_2

Muller, M. J., & Kogan, S. (2010). Grounded Theory Method in HCI and CSCW. *Cambridge: IBM Center for Social Software*, 1–46.

Muller, M., Lange, I., Wang, D., Piorkowski, D., Tsay, J., Liao, Q. V., Dugan, C., & Erickson, T. (2019). How Data Science Workers Work with Data: Discovery, Capture, Curation, Design, Creation. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 126:1--126:15. https://doi.org/10.1145/3290605.3300356

Nardi, B. A., & O'Day, V. L. (1999). Information ecologies: Using technology with heart. *First Monday*. https://doi.org/10.1177/089443930001800113

National Neighborhood Indicators Partnership. (n.d.). *Data and Tech Training Catalog*. https://www.neighborhoodindicators.org/data-tech/training

Nygaard, K., & Terje Bergo, O. (1975). The Trade Unions - New users of research. *Personnel Review*, *4*(2), 5–10. https://doi.org/10.1108/eb055278

O'Connell, K., Lee, Y., Peer, F., Staudaher, S. M., Godwin, A., Madden, M., & Zegura, E. (2016). *Making Public Safety Data Accessible in the Westside Atlanta Data Dashboard*.

O'Reilly, T. (2011). Government as a Platform. *Innovations: Technology, Governance, Globalization*, *6*(1), 13–40. https://doi.org/10.1162/INOV_a_00056

Olivier, P., & Wright, P. (2015). Digital civics: Taking a local turn. In *Interactions*. https://doi.org/10.1145/2776885

Ortner, S. B. (2006). *Anthropology and Social Theory: Culture, Power, and the Acting Subject*. Duke University Press. https://doi.org/10.1177/1368431012463809

Otto, T., & Smith, R. C. (2013). Design anthropology: A distinct style of knowing. In *Design anthropology: Theory and practice* (pp. 1–29).

Parvin, N. (2018). Doing Justice to Stories: On Ethics and Politics of Digital Storytelling.

Engaging Science, Technology, and Society. https://doi.org/10.17351/ests2018.248

Perez, T., & Rushing, R. (2007). *THE CITISTAT MODEL How Data-Driven Government Can Increase Efficiency & Effectiveness*.

Pilemalm, S. (2018). Participatory Design in Emerging Civic Engagement Initiatives in the New Public Sector. *ACM Transactions on Computer-Human Interaction*, *25*(1). https://doi.org/10.1145/3152420

Pilemalm, S., & Timpka, T. (2008). Third generation participatory design in health informatics-Making user participation applicable to large-scale information system projects. *Journal of Biomedical Informatics*. https://doi.org/10.1016/j.jbi.2007.09.004

Pollock, N., & Williams, R. (2009). *Software and organisations: The biography of the enterprise-wide system or how SAP conquered the world*. Routledge. https://doi.org/10.4324/9780203891940

Prost, S., Vlachokyriakos, V., Midgley, J., Heron, G., Meziant, K., & Crivellaro, C. (2019). Infrastructuring food democracy: The formation of a local food hub in the context of socio-economic deprivation. *Proceedings of the ACM on Human-Computer Interaction*. https://doi.org/10.1145/3359159

Puussaar, A., Johnson, I. G., Montague, K., James, P., & Wright, P. (2018). Making open data work for civic advocacy. *Proceedings of the ACM on Human-Computer Interaction*. https://doi.org/10.1145/3274412

Rajapakse, R., Brereton, M., & Sitbon, L. (2018). Design Artefacts to Support People with a Disability to Build Personal Infrastructures. *Proceedings of the 2018 on Designing Interactive Systems Conference 2018 - DIS '18*, 277–288. https://doi.org/10.1145/3196709.3196749

Randall, D. P., Ilana Diamant, E., & Lee, C. P. (2015). Creating sustainable cyberinfrastructures. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/2702123.2702216

Resnick, M., Rusk, N., & Cooke, S. (1998). The computer clubhouse: Technological fluency in the inner city. In W. Schon, D., Sanyal, B. and Mitchell (Ed.), *High Technology and Low-Income Communities. MIT Press. 1998*. MIT Press.

Ribes, D. (2014). Ethnography of scaling, or, how to a fit a national research infrastructure in the room. *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '14*. https://doi.org/10.1145/2531602.2531624

Ribes, D., & Finholt, T. (2009). The Long Now of Technology Infrastructure: Articulating Tensions in Development. *Journal of the Association for Information Systems*. https://doi.org/10.17705/1jais.00199

Rode, J. A. (2011). Reflexivity in digital anthropology. *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems - CHI '11*, 123. https://doi.org/10.1145/1978942.1978961

Rogers, Y. (2011). Interaction design gone wild: Striving for wild theory. *Interactions*. https://doi.org/10.1145/1978822.1978834

Rogers, Y., & Marshall, P. (2017). Research in the Wild. *Synthesis Lectures on Human-Centered Informatics*. https://doi.org/10.2200/s00764ed1v01y201703hci037

Rosner, D. K., Hertz, G., Houston, L., Jackson, S. J., & Rangaswamy, N. (2013). Reclaiming Repair: Maintenance and Mending as Methods for Design. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/2468356.2479674

Rottenburg, R., Merry, S. E., Park, S.-J., & Mugler, J. (2015). *The World of Indicators* (R. Rottenburg, S. E. Merry, S.-J. Park, & J. Mugler (Eds.)). Cambridge University Press. https://doi.org/10.1017/CBO9781316091265

Rouncejield, M., Hughes, J. A., Rodden, T., & Viller, S. (1994). Working with "constant interruption": CSCW and the small office. *ACM Conference on Computer Supported Cooperative Work (CSCW)*. https://doi.org/10.1145/192844.193028

Royster, J. J., Balfour, A., & Burke, C. (2013). *The Georgia Tech/Westside Communities Compact: Ethical Civic Engagement, Global Leadership Development, Sustainable Urban Vitality*.

Sabourin, C. (2016). Responding to the Detroit Water Crisis: The Great Lakes Water Authority and the City of Detroit. *Washington University Journal of Law and Policy*.

Sambasivan, N., & Smyth, T. (2010). The human infrastructure of ICTD. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/2369220.2369258

Sampson, E. E. (1975). On justice as equality. *Journal of Social Issues*, *31*(3), 45–64.

Sandberg, Å. (1979). *"The Demos project: democratic control and planning in working life." In Working on the quality of working life.* (pp. 291–302). Springer.

Sandoval, W. A., & Bell, P. (2004). Design-based Research Methods for Studying Learning in Context: Introduction. *Educational Psychologist*, *39*(4), 199–201. https://doi.org/10.1207/s15326985ep3904

Santos, R. B. (2016). *Crime analysis with crime mapping*. Sage publications.

Sarikaya, A., Correll, M., Bartram, L., Tory, M., & Fisher, D. (2019). What Do We Talk About When We Talk About Dashboards? *IEEE Transactions on Visualization and Computer Graphics*, 682–692.

Sawicki, D. S., & Craig, W. J. (1996). The democratization of data: Bridging the gap for community groups. *Journal of the American Planning Association*. https://doi.org/10.1080/01944369608975715

Sawyer, S., Erickson, I., & Jarrahi, M. H. (2019). Infrastructural Competence. In D. Ribes & J. Vertesi (Eds.), *Digital STS Handbook*. University of Chicago Press. https://doi.org/10.2307/j.ctvc77mp9.22

Schrock, A. R. (2015). Civic hacking as data activism and advocacy: A history from publicity to open government data. *New Media and Society*. https://doi.org/10.1177/1461444816629469

Schrock, A. R. (2018). *Civic tech: Making technology work for people*.

Sein, M. K., Ahmad, I., & Harindranath, G. (2008). Sustaining ICT for development projects: The case of Grameenphone CIC. *Telektronikk*.

Sein, M. K., & Furuholt, B. (2012). Intermediaries: Bridges across the digital divide. *Information Technology for Development*, *18*(4), 332–344.

https://doi.org/10.1080/02681102.2012.667754

Shapiro, D., Hughes, J., Harper, R., Ackroyd, S., & Soothill., K. (1991). Policing information systems: The social context of success and failure in introducing information systems in the police service. *Government and Municipal Information Systems*.

Shen-Hsieh, A., & Schindl, M. (2002). Data visualization for strategic decision making. *Case Studies of the CHI2002|AIGA Experience Design FORUM on - CHI '02*, 1–17. https://doi.org/10.1145/507752.507756

Smith, R. C., & Iversen, O. S. (2018). Participatory design for sustainable social change. *Design Studies*. https://doi.org/10.1016/j.destud.2018.05.005

Smith, R. C., & Kjærsgaard, M. G. (2015). Preface: Design anthropology in participatory design. *Interaction Design and Architecture(s) Journal*, *25*, 73–80.

Star, S. (1999). The ethnography of infrastructure. *American Behavioral Scientist*. https://doi.org/10.1177/00027649921955326

Star, S. (2010). This is Not a Boundary Object: Reflections on the Origin of a Concept. *Science, Technology & Human Values*, *35*(5), 601–617. https://doi.org/10.1177/0162243910377624

Star, S., & Bowker, G. (2002). How to Infrastructure. In L. A. Lievrouw & S. Livingstone (Eds.), *Handbook of New Media: Social Shaping and Consequences of ICTs* (pp. 151–162). SAGE Publications, Ltd. https://doi.org/10.4135/9781848608245.n12

Star, S., Bowker, G., & Neumann, L. (2003). Transparency At Different Levels of Scale: Convergence between Information Artifacts and Communities of Practice. In A. P. Bishop, N. A. Van House, & B. B. Pfeil (Eds.), *Digital Library Use* (pp. 241–269). MIT Press.

Star, S., & Griesemer, J. (1989). Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907–39. *Social Studies of Science*. https://doi.org/10.1177/030631289019003001

Star, S. L. (1985). Scientific Work and Uncertainty. *Social Studies of Science*.

https://doi.org/10.1177/030631285015003001

Star, S. L. (2012). Living Grounded Theory. In A. Bryant & K. Charmaz (Eds.), *The SAGE Handbook of Grounded Theory*. https://doi.org/10.4135/9781848607941.n3

Star, S., & Ruhleder, K. (1996). Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces. *Information Systems Research*, *7*(1), 111–134. https://doi.org/10.1287/isre.7.1.111

Suchman, L. (1987). Plans and Situated Actions. In *Cambridge University Press*. https://doi.org/10.1002/asi.20714

Suchman, L. (2002). Located Accountabilities in Technology Production. *Scandinavian Journal of Information Systemsournal of Information Systems*, *14*(2).

Suchman, L. A. (1983). Office procedure as practical action: Models of work and system design. *ACM Transactions on Information Systems*, *1*(4), 320–328. https://doi.org/10.1145/357442.357445

Suchman, L., & Wynn, E. (1984). PROCEDURES AND PROBLEMS IN THE OFFICE. *Office Technology and People*. https://doi.org/10.1108/eb022630

Tabangin, D. R., Flores, J. C., & Emperador, N. F. (2008). Investigating Crime Hotspot Places and their Implication to Urban Environmental Design : A Geographic Visualization and Data Mining Approach. *World Academy of Science, Engineering and Technology*.

Taylor, A. S., Lindley, S., Regan, T., & Sweeney, D. (2015). Data-in-Place: Thinking through the Relations Between Data and Community. *CHI 2015, Crossings*, 2863–2872. https://doi.org/10.1145/2702123.2702558

Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. *Big Data and Society*. https://doi.org/10.1177/2053951717736335

Taylor, N., Cheverst, K., Wright, P., & Olivier, P. (2013). Leaving the wild: Lessons from community technology handovers. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/2470654.2466206

Trapp, A. C., Teytelboym, A., Martinello, A., Andersson, T., & Ahani, N. (2018). *Placement Optimization in Refugee Resettlement*.

Turkle, S., & Papert, S. (n.d.). Epistemological pluralism and the revaluation of the concrete. *Journal of Mathematical Behavior*. Retrieved March 17, 2015, from http://kvantti.kapsi.fi/Documents/Turkle Papert - Epistemological Pluralism and the Revaluation of the Concrete - 1992.pdf

Tygel, A. F., & Kirsch, R. (2016). Contributions of Paulo Freire for a Critical Data Literacy: a Popular Education Approach. *The Journal of Community Informatics*, *12*(3).

Vallgårda, A., & Fernaeus, Y. (2015). Interaction Design as a Bricolage Practice. *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction - TEI '14*, 173–180. https://doi.org/10.1145/2677199.2680594

van Schalkwyk, F., Willmers, M., & McNaughton, M. (2016). Viscous Open Data: The Roles of Intermediaries in an Open Data Ecosystem. *Information Technology for Development*. https://doi.org/10.1080/02681102.2015.1081868

Vertesi, J. (2014). Seamful Spaces: Heterogeneous Infrastructures in Interaction. *Science Technology and Human Values*, *39*(2), 264–284. https://doi.org/10.1177/0162243913516012

Vertesi, J., & Dourish, P. (2011). The Value of Data : Considering the Context of Production in Data Economies. *Computer Supported Cooperative Work*, 533–542. https://doi.org/10.1145/1958824.1958906

Vitak, J., Liao, Y., Kumar, P., & Subramaniam, M. (2018). Librarians as information intermediaries: Navigating tensions between being helpful and being liable. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-319-78105-1_80

Vlachokyriakos, Vasilis, Crivellaro, C., Le Dantec, C. A., Gordon, E., Wright, P., & Olivier, P. (2016). Digital Civics: Citizen Empowerment With and Through Technology. *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. https://doi.org/10.1145/2851581.2886436

Vlachokyriakos, Vasilis, Crivellaro, C., Wright, P., & Olivier, P. (2018a).

Infrastructuring the Solidarity economy: Unpacking strategies and tactics in designing social innovation. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3173574.3174055

Vlachokyriakos, Vasilis, Crivellaro, C., Wright, P., & Olivier, P. (2018b). Infrastructuring the Solidarity economy: Unpacking strategies and tactics in designing social innovation. *Conference on Human Factors in Computing Systems - Proceedings*, *2018-April*. https://doi.org/10.1145/3173574.3174055

Vlachokyriakos, Vasillis, Comber, R., Crivellaro, C., Taylor, N., Kuznetsov, S., Kavanaugh, A., Le Dantec, C. A., & Kim, B. J. (2015). Designing Alternative Systems for Local Communities. *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '15*, 2333–2336. https://doi.org/10.1145/2702613.2702657

Wagstaff, G. F. (1994). Equity, equality, and need: Three principles of justice or one? An analysis of "equity as desert." *Current Psychology*, *13*(2), 138–152. https://doi.org/10.1007/BF02686797

Walker, D. (2018). "Data justice" by design: Building engagement through civic technologies. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3170427.3173024

Wheeler, E., & Dillahunt, T. R. (2018). Navigating the job search as a low-resourced job seeker. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3173574.3173622

Whitney, K. (2011). *The Bluff: Atlanta's forgotten community*. 11alive.Com. http://downtown.11alive.com/news/news/88834-bluff-atlantas-forgotten-community

Williams, R., & Pollock, N. (2012). Moving beyond the single site implementation study: How (and why) we should study the biography of packaged enterprise solutions. *Information Systems Research*. https://doi.org/10.1287/isre.1110.0352

Wolff, A., Gooch, D., Cavero Montaner, J. J., Rashid, U., & Kortuem, G. (2016). Creating an Understanding of Data Literacy for a Data-driven Society. *The Journal of Community Informatics*.

Wood, C., Poslad, S., Kaniadakis, A., & Gabrys, J. (2017). What Lies Above: Alternative User Experiences Produced Through Focussing Attention on GNSS Infrastructure.

Proceedings of the 2017 Conference on Designing Interactive Systems - DIS '17, 161–172. https://doi.org/10.1145/3064663.3064757

Wynn, E. (1979). *Office conversation as an information medium*. University of California, Berkeley.